CN-124,085

STATISTICAL STUDY OF A UNIFORM AND

A NORMAL RANDOM NUMBER GENERATOR

By

George Chris Canavos

Thesis submitted to the Graduate Faculty of the

Virginia Polytechnic Institute

in candidacy for the degree of

MASTER OF SCIENCE

in

Statistics

December 1966

STATISTICAL STUDY OF A UNIFORM AND

A NORMAL RANDOM NUMBER GENERATOR

by

George Chris Canavos

Thesis submitted to the Graduate Faculty of the

Virginia Polytechnic Institute

in candidacy for the degree of

MASTER OF SCIENCE

in

Statistics

APPROVED

Chairman, Richard G. Krutchkoff

Boyd Harshbarger

Raymond H. Myers

Whitney L. Johnson

December 1966

Blacksburg, Virginia

TABLE OF CONTENTS

I. INTRODUCTION

1.1 Objectives

This thesis will present a detailed statistical study of a uniform and a normal random number generator as well as the implementation of each on the IBM-7094 digital computer.

Basically, this thesis will attempt to discover desirable or undesirable features of such generators by subjecting them to a fairly extensive set of statistical tests. Based on the results of these tests, one should have a fairly good indication concerning such important aspects as distribution properties and the degree of randomness exhibited by the generators. However, it should be pointed out that no matter how strongly the results may indicate a favorable conclusion, one should exercise some degree of caution not to be overly optimistic. The reason for this is the fact that the power of some statistical tests of hypothesis considered in this study cannot be determined.

A computer program has been written by the author to generate the necessary pseudo-random numbers for the study. The computer program attempts to combine the basic advantages of FORTRAN IV and MAP (MACRO ASSEMBLY PROGRAM) computer languages so that optimum efficiency can be attained in the generation of pseudo-random numbers.

1.2 Background

In order to be able to solve practical problems in which the construction of some random process is required, it becomes necessary to have some random number generator so that the random process may be simulated. The solutions to such practical problems depends, of course, on how good the given pseudo-random number generator really is. Two of the most widely used generators are those that generate uniformly distributed pseudo-random quantities and normally distributed pseudo-random quantities. The former has the attractive feature of being used to obtain random events obeying other distribution functions.

Since high-speed digital computers have become easily accessible, it is desirable and advantageous to generate by some deterministic means a sequence of pseudo-random numbers obeying some specified distribution function. Such sequences may appear to be random even if, upon closer and longer observation, certain patterns become evident. Therefore, the essence of such a task is to find a generator that exhibits as few patterns as possible and still maintain its statistical behavior according to the specified distribution function. It is important to define a sequence of random numbers generated by some deterministic means as a sequence of pseudo-random numbers because the word "random" no longer maintains its basic meaning in its entirety. Henceforth, whenever the word "random" appears in this thesis it shall imply pseudo-random.

Many attempts have been made, with varying degrees of success, to contribute to the state-of-the-art of random number generators. Most of the existing uniform random number generators are of the form

$$X_{i+1} = aX_i + b \pmod{M} \qquad i = 1, 2, \ldots \tag{1.2.1}$$

with the modulo depending on whether a given computer is binary or decimal. A significant drawback to these is the requirement of extensive testing by trial and error to determine the constants a and b that appear to produce the most satisfactory results. In many papers that have been written, the determination of a and b overshadows the investigation. See for example reference 20.

It is the purpose of this thesis to use a significantly different uniform random number generator in which the parameters are specified conclusively by the theory of the generator. As a result, no additional effort is needed to produce optimal conditions. In addition, the normal random number generator is based on a direct transformation from the uniform to the normal distribution thus eliminating approximate techniques such as that by Hastings.[1] Briefly, the Hastings technique considers the integral

$$u = \int_{X(u)}^{\infty} e^{-\frac{1}{2}t^2}\, dt \qquad 0 < u \leq 0.50 \tag{1.2.2}$$

and approximates $X(u)$ where $X \sim N(0,1)$ by the polynomial relation

$$X(u) = \eta - \left\{\frac{2.30753 + 0.27061\eta}{1 + 0.99229\eta + 0.04481\eta^2}\right\} \tag{1.2.3}$$

where $\eta = \sqrt{\ln \frac{1}{u^2}}$.

[1] Hastings, C., Jr.: Approximations for Digital Computers, Princeton University Press, Princeton, New Jersey, 1955, p. 191.

II. THE GENERATORS

This chapter contains a mathematical presentation of the uniform and normal random number generators considered for this study. The uniform random number generator is based on a sequence of zeros and ones generated by an n^{th} degree maximal length linear recursion relation. This relation is associated with a primitive polynomial with coefficients whose values are either zero or one and a degree n equal to the word length of a given digital computer. As stated previously, the computer toward which this study is oriented is the IBM-7094 which is a binary machine; thus the choice of the (0,1) sequence.

The normal random number generator is based on a direct transformation from two random variables distributed uniformly on the unit interval (0,1).

2.1 The Uniform Random Number Generator

Let $a = \{a_k\}$ be the sequence of zeros and ones generated by the linear recursion relation

$$a_k = c_1 a_{k-1} + c_2 a_{k-2} + \cdot \cdot \cdot + c_n a_{k-n} \pmod 2 \qquad k = 1, 2, \ldots \tag{2.1.1}$$

for any given set of integers c_i $(i = 1, 2, \ldots, n)$ each having the value of zero or one where c_i $(i = 1, 2, \ldots, n)$ are the coefficients of some polynomial

$$f(X) = 1 + c_1 X + c_2 X^2 + \cdot \cdot \cdot + X^n \tag{2.1.2}$$

Choose $f(X)$ so that it is primitive over the Galois field of order two - GF(2). (One defines a Galois field as any finite field; hence, GF(2) is that finite field containing the elements zero and one.) Let $c_n = 1$ and say that $f(X)$ is a primitive n^{th} degree polynomial over GF(2). (A primitive n^{th} degree polynomial over GF(2) may be defined as that polynomial whose roots are primitive $(2^n-1)^{th}$ roots of unity.) As a result, $\{a_k\}$ is said to be a maximal length linearly recurring sequence modulo 2 (ref. 1).

From (2.1.1), notice that a_k is determined solely by the n-tuple $(a_{k-1}, a_{k-2}, \ldots, a_{k-n})$ of terms preceding it. Similarly, a_{k+1} is a function solely of $(a_k, a_{k-1}, \ldots, a_{k-n+1})$. As a result, each such n-tuple has a unique successor governed by the recursion relation (2.1.1).

Let p be the period of the linear recurring sequence $\{a_k\}$. The period of $\{a_k\}$ has to be the same as the period with which an n-tuple repeats. Obviously, p cannot be greater than 2^n-1, where n is the degree of the polynomial $f(X)$, because the n-tuple $(0,0,\ldots,0)$ is always followed by $(0,0,\ldots,0)$. It has been shown (refs. 2 and 3) that a necessary and sufficient condition that $p = 2^n - 1$ is that $f(X)$ be primitive n^{th} degree polynomial over GF(2). Such linearly recurring sequences have been extensively studied and used as codes in communication theory (ref. 4).

Two properties of interest are the following (ref. 2):

$$(1) \quad \sum_{k=1}^{p} a_k = \frac{p+1}{2} = 2^{n-1}$$

(2) Every nonzero binary n-vector $(b_1, b_2, \ldots, b_n)$ occurs exactly once per period as n consecutive binary digits in $\{a_k\}$ where n is the degree of $f(X)$.

As a simple example to illustrate the preceding, let $f(X) = 1 + X + X^2$ be the primitive polynomial of degree $n = 2$ over GF(2) (ref. 5); then $c_1 = c_2 = 1$. Clearly, the linear recursion relation is

$$a_k = a_{k-1} + a_{k-2} \pmod 2 \tag{2.1.3}$$

The period $p = 2^2 - 1 = 3$ indicates that $\{a_k\}$ should repeat after its first three elements have been determined. For every such sequence, let $a_o = 1$ - if $a_o = 0$ all successive a_k would also be zero - then by (2.13)

$$\left.\begin{aligned} a_1 &= a_o = 1 \\ a_2 &= a_1 + a_o = 0 \\ a_3 &= a_2 + a_1 = 1 \end{aligned}\right\} \text{the period of the sequence}$$

$$\begin{aligned} a_4 &= a_3 + a_2 = 1 \\ a_5 &= a_4 + a_3 = 0 \\ a_6 &= a_5 + a_4 = 1 \end{aligned}$$

.
.
.
etc.

Notice that the elements of $\{a_k\}$ repeat themselves after the first three have been determined. Property (1) is satisfied in that $\sum_{k=1}^{3} a_k = a_1 + a_2 + a_3 = 2$. Property (2) is also satisfied since one has binary vectors (1,0,1), (1,1,0), (0,1,1) that occur exactly once per period.

More generally, let $f(X)$ and $\{a_k\}$ be defined. Consider the sequence of numbers of the form

$$Y_k = \sum_{t=1}^{L} 2^{-t} a_{qk+r-t} \tag{2.1.4}$$

where the optimal value of L is equal to n, the degree of $f(X)$; r is an arbitrarily chosen integer, $0 \leq r \leq 2^n - 1$; and q is any integer greater or equal to L chosen so that q and $p = 2^n - 1$ are relatively prime. Clearly, Y_k is the binary expansion of a number whose binary representation is L consecutive digits in a, and each Y_k is spaced q digits apart. From (2.1.4), it can be seen that such numbers always lie in the interval $0 < Y_k < 1$.

Equation (2.1.4) is due to R. C. Tausworthe (ref. 1) who has shown that such sequence of numbers are uniformly distributed on the unit inverval (0,1) with mean $\mu = \frac{1 - 2^{-L}}{2}$ and variance $\sigma^2 = \frac{1}{12}$; hence for large L, the mean and variance of Y_k are identically the same as the mean and variance of the uniform distribution function in the unit interval. As a result, equation (2.1.4) represents the uniform random number generator that is implemented for this study.

A primitive polynomial of degree 35 which is equal to the word length of the IBM-7094 was obtained from a table of primitive polynomials over GF(2) published by Watson (ref. 5). The polynomial

$$f(X) = X^{35} + X^2 + 1 \tag{2.1.5}$$

is associated with the linearly recurring sequence

$$a_k = a_{k-2} + a_{k-35} \pmod 2 \tag{2.1.6}$$

for $k \geq 35$; for $1 \leq k \leq 34$

$$a_k = 1 \text{ if } k \text{ is even}$$

$$a_k = 0 \text{ if } k \text{ is odd}$$

For this study, L and q were set equal to 35 which is relatively prime to the period $p = 2^{35} - 1$. Therefore, one can produce, by equation (2.1.4), precisely $2^{35} - 1$ Y's before repetition occurs. Of course, this study considers only a small portion of these, say $N = 10{,}000$, and attempts to discover their various properties as will be seen later.

2.2 The Normal Random Number Generator

G. E. P. Box and Mervin E. Muller (ref. 6) have derived a method to obtain a pair of independent random variables normally distributed with mean zero and variance one from two independent random variables from the same uniform distribution on the interval (0,1). A brief presentation of their approach follows.

Let U_1, U_2 be independent random variables uniformly distributed on the unit interval. Consider the random variables

$$X_1 = (-2 \log_e U_1)^{1/2} \sin 2\pi U_2 \tag{2.2.1}$$

$$X_2 = (-2 \log_e U_1)^{1/2} \cos 2\pi U_2 \tag{2.2.2}$$

By giving attention to principal values, obtain the inverse relationships

$$U_1 = e^{-(X_1^2+X_2^2)/2} \tag{2.2.3}$$

$$U_2 = \frac{1}{2\pi} \arctan\left(\frac{X_2}{X_1}\right) \tag{2.2.4}$$

Then, the joint density of X_1, X_2 is

$$f(X_1,X_2) = \frac{1}{2\pi} e^{-(X_1^2+X_2^2)/2} \tag{2.2.5}$$

but

$$f(X_1,X_2) = \frac{1}{\sqrt{2\pi}} e^{-X_1^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-X_1^2/2} = f(X_1)f(X_2) \tag{2.2.6}$$

Hence, X_1, X_2 are a pair of independent random variables from the same normal distribution with mean zero and unit variance.

Box and Muller's approach is based on the following considerations: the probability density $f(X_1, X_2)$ is constant on circles; hence $\theta = \arctan(X_2/X_1)$ is uniformly distributed on the interval $(0,2\pi)$. Further, the square of the length of radius vector $r^2 = X_1^2 + X_2^2$ has a chi-squared distribution with two degrees of freedom. Since U_1 is uniformly distributed in the interval $(0,1)$, then $-2 \log_e U_1$ is distributed as a chi-squared with two degrees of freedom. By proceeding in the reverse order, Box and Muller arrive at (2.2.1) and (2.2.2).

Since equations (2.2.1) and (2.2.2) contain square roots, trigonometric functions, and natural logarithms, their accuracy depends, in part, on the accuracy of the available library programs which compute these functions. The IBM-7094 library programs possess fairly high accuracy; therefore, for the major part, the accuracy with which equations (2.2.1) and (2.2.2) are computed depends on the degree of uniformity of U_1 and U_2.

III. THE IMPLEMENTATION OF THE GENERATORS ON THE IBM-7094 COMPUTER

Needless to say, it would be a colossal task, if not impossible, to attempt to generate uniform or normal random numbers by the methods discussed in Chapter II without the use of some high-speed digital computer. For this reason, a computer program has been written by the author to facilitate this study.

The program was written by considering the most desirable features of the IBM-7094 and its internal language so that a random number may be generated with a minimum amount of computer time. A listing of the computer program appears along with a detailed discussion of its contents.

3.1 Computer Program

Basically, the computer program is written to provide two options: (a) the generation of one random number uniformly distributed on the interval (0,1) according to equation (2.1.4), and (b) the generation of one random number normally distributed with zero mean and unit variance according to equation (2.2.1), and, as a consequence, two random numbers uniformly distributed on the unit interval. It is implicit here that the first option be chosen whenever one needs only uniformly distributed random numbers, while the second be chosen whenever one's primary interest is to obtain normally distributed random numbers.

For ease of discussion, the program is divided into four sections. The first two sections are written in Fortran IV using integer arithmetic. Mainly, these two sections generate the sequence a_k of zeros and ones

```
C     A COMPUTER PROGRAM FOR THE GENERATION OF NORMAL AND/OR UNIFORM
C                         PSEUDO-RANDOM NUMBERS
      SUBROUTINE GETRAN(IR,N,L,RN,Y1,Y2)
      COMMON IA(35,2)
      DIMENSION IA2(35,2),IA3(35,2),IA5(35,2),IA1(35,2),IR(2)
      CON=6.2831853
      IF(N.GT.1)GOTO22
C     SECTION 1  INITIALIZATION
      DO2M=1,L
      DO1I=2,34,2
C     STORING A 1 IN A(2),A(4),....,A(34) AND A 0 IN A(1),A(3),....,A(33)
    1 IA2(I,M)=1
      DO9I=1,33,2
    9 IA2(I,M)=0
C     STORING A 1 IN A(35)
    2 IA2(35,M)=1
      N1=35
      DO21MI=1,L
C     LOCATING THE FIRST USABLE 35 ELEMENTS OF SEQUENCE A BASED ON THE
C     ARBITRARILY CHOSEN INPUT INTEGER IR
   18 DO17I=1,35
      IF(I.EQ.1)ITEMP=IA2(34,MI)
      IF(I.EQ.2)ITEMP=IA2(35,MI)
      IF(I.GT.2)ITEMP=IA3(I-2,MI)
```

```
      IA3(I,MI)=ITEMP+IA2(I,MI)
      IF(IA3(I,MI).EQ.2)IA3(I,MI)=0
      IA1(I,MI)=IA2(I,MI)
   17 IA2(I,MI)=IA3(I,MI)
      MAX=35+IR(MI)-1
      N1=N1+35
      IF(N1.GE.MAX)GOTO3
      GOTO18
    3 N2=N1-MAX
      N3=35-N2
      IZ=0
      N4=N3+1
      IF(N2.EQ.0)GOTO33
      DO19I=N4,35
      IZ=IZ+1
   19 IA(IZ,MI)=IA1(I,MI)
   33 CONTINUE
      IF(N3.EQ.0)GOTO21
      DO49I=1,N3
      IZ=IZ+1
   49 IA(IZ,MI)=IA3(I,MI)
   21 N1=35
C     THE FIRST 35 ELEMENTS OF A HAVE BEEN LOCATED AND STORED IN
C     ARRAY IA
      GOTO25
```

```
C       SECTION 2  COMPUTATION OF SUCCESSIVE ELEMENTS OF ARRAY IA IN SETS
C       OF 35 FOR EACH RANDOM NUMBER
     22 DO26M=1,L
        DO26I=1,35
        IF(I.EQ.1)ITEMP=IA(34,M)
        IF(I.EQ.2)ITEMP=IA(35,M)
        IF(I.GT.2)ITEMP=IA5(I-2,M)
        IA5(I,M)=ITEMP+IA(I,M)
        IF(IA5(I,M).EQ.2)IA5(I,M)=0
     26 IA(I,M)=IA5(I,M)
     25 CALL FAITH(L,Y1,Y2)
        IF(L.GT.1)RN=SQRT(-2.*ALOG(Y1))*SIN(CON*Y2)
        RETURN
        END
$IBMAP HOPE
        ENTRY   FAITH
FAITH   SAVE    1
*       INITIALIZATION
        STZ     Y1
        STZ     Y2
        CLA*    3,4
        SUB     =1
        TZE     HERE
*       IF USER NEEDS ONLY UNIFORM RANDOM NUMBERS THE PROGRAM TRANSFERS
*       TO SECTION 4. IF USER NEEDS UNIFORM AND NORMAL RANDOM NUMBERS THE
*       PROGRAM CONTINUES
```

```
*       SECTION 3  THE COMPUTATION OF TWO UNIFORM RANDOM NUMBERS
        CAL     TWOJ
        AXT     1,1
LOOP    ARS     1
        SLW     WRK
        NZT     IA+35,1
        TRA     *+3
        ADD     Y1
        STO     Y1
        CAL     WRK
        NZT     IA+70,1
        TRA     *+3
        ADD     Y2
        STO     Y2
        CAL     WRK
        TXI     *+1,1,1
        TXL     LOOP,1,35
        CLA     Y1
        ARS     8
        ORA     =181
        FAD     =0.
        STO     Y1
        STO*    4,4
        CLA     Y2
        ARS     8
```

```
Y1      BSS      1
Y2      BSS      1
WRK     BSS      1
        CONTRL   //
IA      COMMON   70
        END
```

according to the linear recursion relation (2.1.6) and the given polynomial (2.1.5). The remaining two sections are written entirely in MAP, the symbolic language of the IBM-7094. Section 3 generates two uniform random numbers in the unit interval and serves as an intermediate step toward the determination of one normally distributed random number by equation (2.2.1). If the first option is chosen, Section 3 is omitted and its basic function is assumed by Section 4 where one uniform random number is generated.

The calling sequence of the computer program

SUBROUTINE GETRAN(IR,N,L,RN,Y1,Y2)

is the entrance point to the program. The calling sequence contains information for the necessary input and expected output of the program. The input of the program is comprised of the argument IR, N, and L. The argument IR - corresponding to the integer r of Chapter II - is a one-dimensional array of two different positive fixed integers arbitrarily chosen. For all practical purposes, their values could be anywhere from, say, 50 to 200,000. The argument N controls the specific point of entrance to the computer program. For any one of the two available options, if multiple random numbers are needed, it is essential that the initialization part (Section 1) is omitted after the first time the computer program has been called. Hence, for the first call of the program the argument N has to be equal to the fixed integer one. Otherwise, N may be equal to any fixed integer greater than one. The argument L controls the choice of the two options. If the first option is chosen, L has to be equal to the fixed integer one. If the second option is chosen, the argument L must be equal to the fixed integer two.

Arguments RN, Y1, and Y2 constitute the output of the computer program. For the first option, Y1 will be the location where the one uniform random number will be stored. For the second option, arguments RN, Y1, and Y2 will contain the one normal and the two uniform random numbers, respectively. It is worth noting that all arguments in the calling sequence are dummy arguments and may be named differently. Moreover, the array IA (35,2) must be placed in COMMON by the user.

As an example of the use of the computer program, consider the need for 100 random numbers normally distributed with zero mean and unit variance. Define an array A in which these random numbers will be stored. Then, one way to obtain these random numbers by using SUBROUTINE GETRAN would be as follows:

```
      COMMON IA (35,2)
      DIMENSION MN (2), A (100)
      MN (1) = 3063
      MN (2) = 10275
      DO 1 I = 1,100
      CALL GETRAN (MN, I, 2, X, Y, Z)
    1 A (I) = X
      .
      .
      .
      STOP
      END
```

Notice that the arguments of the calling sequence have been renamed; that is, MN, I, 2, X, Y, Z correspond to IR, N, L, RN, Y1, Y2, respectively. By replacing the argument N, the running subscript I of the DO loop satisfies the requirement that its value be equal to one for the first call of SUBROUTINE GETRAN and greater than one for all subsequent calls.

It has been noted briefly that Sections 1 and 2 generate the sequence a_k of zeros and ones. More specifically, the primary purpose of Section 1 is the initialization of the computer program. In this section, the first usable 35 elements of $\{a_k\}$ necessary to compute one uniform number according to equation (2.1.4) are located and determined. Their location depends strictly on the value of the arbitrarily chosen integer stored in IR(1). Recall that the subscript of $\{a_k\}$ in

$$Y_m = \sum_{t=1}^{L} 2^{-t} a_{qm+r-t}$$

depends on the value of the integer r (IR) and the integer $q = 35$. Hence for $m = 1$, the first element of $\{a_k\}$ needed in the sum is clearly a_{35+r-1}, while the last is $a_{35+r-35}$ where $L = 35$. The location of these 35 elements of $\{a_k\}$ implies the immediate calculation of the elements of $\{a_k\}$ for all k up to and including the element a_{35+r-1} according to the linear recursion relation (2.1.6). If the second option is chosen, the procedure is repeated using the second arbitrarily chosen integer stored in IR(2). Otherwise, the task of Section 1 is completed. For any value of the argument N greater than one, the computer program will assume that initialization has occurred; thus, it will omit Section 1 and proceed to Section 2.

It is reasonable to assume that one would want to call the program more than once using the same input, except for the value of the argument N, in order to obtain multiple random numbers based on the option chosen. Hence for each call of the computer program at least 35 more elements of $\{a_k\}$ have to be determined. This becomes the task of Section 2. Recall that the first set of 35 elements of $\{a_k\}$ with respect to the integer IR(1) were located and determined in Section 1. (Of course, the procedure is quite the same for the integer IR(2), if it is required by the option.) Therefore, to compute an additional 35 elements, consider

$$Y_m = \sum_{t=1}^{35} 2^{-t} a_{qm+r-t}$$

when $m = 2$. Clearly, the first element needed in this sum is a_{70+r-1} while the last is $a_{70+r-35}$. Notice that $a_{70+r-35}$ is that element of $\{a_k\}$ immediately following a_{35+r-1}, the last element of the first set. Hence, for $m = 2$, the 35 elements of $\{a_k\}$ needed for the sum are those that follow immediately after the first set of 35 elements determined in Section 1. Therefore, for any m it is only necessary to compute 35 elements of $\{a_k\}$ based on the preceding set as specified by the theory of the generator in Chapter II. After initialization, Section 2 assumes this task.

Recall that Section 3 is associated with the second option, and Section 4 implies the first option. Entrance to either one of these sections implies that the necessary elements of $\{a_k\}$ have been determined and stored. Because of the similarity in the purpose of these two sections, it is only necessary to discuss Section 4 and imply the procedure for Section 3.

Section 4 is programed in MAP. The advantages for the choice of this computer language become evident by the following discussion. The IBM-7094 numerical register contains 36 bit positions, the first of these being the sign position. Consider storing in the 36 bit positions the binary number 1 followed by 35 zeros (in octal this number is written as 400000000000). Thus, the configuration of the numerical register would appear as follows:

S	1	2																																34	35
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Shift the contents of the numerical register one place to the right. As a result, the number 1 appears under bit position one, and the 0 that was under bit position 35 is shifted in the sign position. Choose the binary point to be to the left of bit position one. Thus, bit position one has the value of 2^{-1} since it contains the number 1. As a matter of fact, the numerical register has a value of 2^{-1} since all other bit positions contain zeros. Similarly, if bit position two contained the number 1 with zeros in all other bit positions, the numerical register would have the value of 2^{-2}. The argument can be extended to include every bit position of the numerical register. Therefore, one may think of each bit position of the register as being capable of representing successive negative powers of the base 2 with the power being equal to the bit position.

Let the value of the numerical register be equal to 2^{-1}. Consider, as an example, some 35 elements of $\{a_k\}$ to be the following:

$$a_{35m+r-t} = \{1,0,1,1,0,0,0,1,1,1,0,0,1,1,0,0,0,0,1,0,1,1,0,0,0,1,1,1,0,0,1,1,0,0,1\}$$

for $t = 1, 2, \ldots, 35$ and some m. Consider a way to test each of these 35 elements sequentially beginning with the first to determine whether its value is zero or one. If it is one, the value of the register, presently equal to 2^{-1} since bit position one contains the number one, is added to a quantity Y1 which initially was set equal to zero. If the element being tested is zero, nothing is added to Y1. Hence, for the example, Y1 is equal to 2^{-1} because the first element is one. Proceed by shifting the contents of the numerical register one place to the right. Now bit position two contains the number 1 with zeros in all other bit positions. This implies that the value of the register is equal to 2^{-2}. Test the next consecutive element. Add the contents of the register to Y1 only if the value of the element is one. Otherwise, Y1 remains the same. For the example, Y1 remains equal to 2^{-1} because the second element is zero. Again shift the contents of the numerical register one place to the right. As a result, the number 1 appears under bit position three, and the value of the register is 2^{-3}. Test the third element. If one, add the value of the register to Y1; otherwise, Y1 remains the same. For the example, the third element is one; therefore, Y1 is now equal to $2^{-1} + 2^{-3}$. Continue this procedure until all 35 elements have been tested sequentially. At the end, Y1 should represent the value of one random number uniformly distributed on the interval (0,1). For the example, Y1 would be equal to

$$2^{-1} + 2^{-3} + 2^{-4} + 2^{-8} + 2^{-9} + 2^{-10} + 2^{-13} + 2^{-14} + 2^{-19} + 2^{-21} + 2^{-26} + \\ + 2^{-27} + 2^{-28} + 2^{-31} + 2^{-32} + 2^{-35}$$

based on the seventeen nonzero elements and the corresponding contents of the numerical register.

The preceding approach offers the attractive advantage of considerable reduction of computer time. Recall that the mathematical definition of the uniform number

$$Y_m = \sum_{t=1}^{L} 2^{-t} a_{am+r-t}$$

implies the multiplication of each element of $\{a_k\}$ needed in the sum by the number 2 which has been raised to some negative power equal to the value of the subscript t. The raising of a given number to a power, division, and multiplication, require considerably more time in most digital computers than addition or subtraction. Notice that all elements of $\{a_k\}$ are either zero or one. Therefore the product

$$2^{-t} a_{qm+r-t}$$

really implies add or not add 2^{-t} to the sum depending on whether the corresponding element of $\{a_k\}$ is one or zero. Since the numerical register of the IBM-7094 was so constructed to produce successive 2^{-t} for $t = 1, 2, \ldots, 35$, the operations needed to compute

$$2^{-t} a_{qm+r-t}$$

have been reduced to one; namely, addition. FORTRAN IV does not have the necessary flexibility to permit this approach; thus the choice of MAP. From the tests that have been made, it has been determined that the amount of computer time necessary to determine one random number by this method is approximately 0.0015 second.

IV. TESTS FOR RANDOMNESS

If the formulae discussed in Chapter II are to be useful in simulating random events or processes, it is essential that each generated sequence of numbers possess desirable features of randomness.

Intuitively speaking, a random sequence of numbers is a sequence in which the specific values of the elements are not at all a function of their position in the sequence. In other words, any particular order the values of some random sequence present themselves is no more likely to occur than any other ordering.

For the overall investigation, a total of two-hundred (200) sequences each containing 10,000 elements were generated. Of these, one-hundred were generated using the uniform number generator (2.1.4), fifty were generated using equation (2.2.1), and the remaining fifty were generated using equation (2.2.2). Equations (2.2.1) and (2.2.2) comprise the normal random number generator.

Three tests have been performed on all sequences for the detection of nonrandomness. Two of these are runs tests, and the third is a test on serial correlation. These tests are intended to serve a dual purpose: (a) to detect nonrandomness, and (b) to create a system of checks between the results of one test as compared with the results of another. All the necessary computations were performed on the IBM-7094 with the results appearing at the end of this chapter.

4.1 Runs Test

To detect a lack of randomness, one must decide whether a given sequence of numbers generated by such deterministic means as those of

Chapter II is likely to occur by chance or some assignable causes are indicated. A technique with which such a decision can be made is based on the order in which the particular values of a sequence of numbers were obtained. Furthermore, this order depends on the number of runs exhibited in the sequence.

Given a sequence of numbers, consider the assignment of all elements in the sequence into two classes, A and B, class A containing all elements in the sequence that are greater than the mean while class B contains all elements that are less than or equal to the mean of a specified distribution function. Let n_A, n_B be the total number of elements observed in class A and class B, respectively. Then, a run is defined as a succession of elements from the same class contained between elements of a different class. Hence, for a given sequence of numbers, the total number of runs is always one plus the number of unlike neighbors in the given sequence. For example, the sequence

a a b a b b a b a a a

has six unlike neighbors and, therefore, seven total number of runs.

Let r denote the total number of runs contained in a given sequence of numbers. Since r can take on any value within some domain, then r is a random variable that gives an indication on whether a sequence of numbers may be looked upon as random. This can be illustrated by considering the following example: Suppose one tosses a coin fifty times resulting in a sequence of only two runs consisting of twenty-five heads followed by twenty-five tails. He would, of course, strongly suspect that the probability of success had not been the same from trial to trial. If, on the other hand, the sequence of fifty tosses contained fifty runs consisting of alternating heads and tails, the suspicion would be that the trials had not been independent.

Consider the question of testing the null hypothesis that a given sequence of numbers appears to be random. The argument applied to the example can also be applied to test this null hypothesis based on total number of runs. It is apparent that if a given sequence of numbers is random, the elements of class A or B should be well mixed and r should neither be too small nor too large. If there exist long successions of elements of the same class followed by long successions of elements of the other class, it would be reasonable to conclude some biased departure from the true probability structure which would tend to reduce r. The other extreme would be when the elements from both classes are alternating with a very high frequency. In this case, the apparent conclusion would be that the value of an element depends on the value of the preceding one. This, of course, would tend to make r fairly large. Therefore, the test is performed by counting the total number of runs in a given sequence, accepting the null hypothesis if for some specified number r_o and r_1, $r_o \leq r \leq r_1$, and rejecting it otherwise.

In order that one may specify r_o and r_1 for a given level of significance, the distribution of the random variable r is needed. It has been shown (refs. 7 and 12) that the density of r is given by

$$P(r) = \frac{2\binom{n_A - 1}{k - 1}\binom{n_B - 1}{k - 1}}{\binom{n_A + n_B}{n_A}} \qquad k = r/2 \qquad (4.1.1)$$

if r is even and

$$P(r) = \frac{\binom{n_A - 1}{k}\binom{n_B - 1}{k - 1} + \binom{n_A - 1}{k - 1}\binom{n_B - 1}{k}}{\binom{n_A + n_B}{n_A}} \qquad k = \frac{r - 1}{2} \tag{4.1.2}$$

if r is odd. Hence to test the null hypothesis in question with a probability α for the Type I error, one finds integers r_o and r_1 so that as nearly as possible

$$\sum_{r=0}^{r_o} P(r) = \frac{\alpha}{2} \tag{4.1.3}$$

and

$$\sum_{r=0}^{r_1} P(r) = 1 - \frac{\alpha}{2} \tag{4.1.4}$$

and rejects the null hypothesis if the observed r is either less than r_o or greater than r_1. It is apparent that the computations involved in (4.1.1) or (4.1.2) are quite lengthy especially if n_A and n_B are fairly large. However, it is believed that if both n_A and n_B are larger than 10, the distribution of r becomes approximately normal (refs. 7 and 17) with a mean

$$\mu_r = \frac{2n_A n_B}{n_A + n_B} + 1 \qquad n_A + n_B = N \tag{4.1.4}$$

and a variance

$$\sigma_r^2 = \frac{2n_A n_B(2n_A n_B - n_A - n_B)}{(n_A + n_B)^2(n_A + n_B - 1)} \tag{4.1.5}$$

The approximation is improved as n_A and n_B become large. For $N = 10{,}000$, it is reasonable to expect the approximation to be fairly good since n_A and n_B will be quite larger than 10. Thus to test the null hypothesis in question, consider the statistic

$$Z = \frac{r - \mu_r}{\sigma_r} \qquad (4.1.6)$$

where Z is the value of a random variable having the standard normal distribution, and r is the observed number of runs of a given sequence. For some probability α for Type I error, the null hypothesis is rejected if the computed Z is less than $Z_{1/2}\alpha$ or if it is greater than $Z_{1-\frac{1}{2}}\alpha$ of the standard normal distribution. Rejection implies that the given sequence may be declared nonrandom based on the test on total number of runs. This test has been performed on each of the 200 sequences generated for this study with α equal to 0.05. The results, found in tables (4.10), (4.11), and (4.12), indicate the following:

(a) For the one-hundred sequences generated by equation (2.1.4), four sequences were rejected and declared nonrandom. By this result, there is an indication that the uniform number generator (2.1.4) generates sequences that appear to be random based on the observable number of total runs.

(b) For the fifty sequences generated by equation (2.2.1), three sequences were rejected and declared nonrandom. This indicates that equation (2.2.1) generates sequences that appear to be random based on the observable number of total runs.

(c) For the fifty sequences generated by equation (2.2.2), there were no rejections. The indication is the same as in (a) and (b); however, further testing, as will be discussed in the following section, is recommended.

One should note here that the preceding test based on total number of runs is somewhat poor because it is effective in detecting nonrandomness only when a given sequence contains too many or too few runs. Situations could arise when a given sequence produces the correct number of total runs, but contains serious types of nonrandomness. For this reason, a test based on counting the number of runs of various lengths will be discussed in the following section.

4.2 Runs of Various Lengths

It was pointed out in the previous section that a test based on the total number of runs in a sequence of numbers can be deceiving. However, a test based on counting the number of runs of various lengths is less likely to be deceived because the observable number of total runs is subdivided into runs of various lengths that can be easily compared to their corresponding theoretical expectations. A run above or below the mean of the specified distribution function is defined as follows: If l successive elements of a given sequence of numbers are greater (or less) than the mean, and both the preceding and following elements are less (or greater) than the mean, this is recorded as a run of length l above (or below) the mean.

Let $r_{A,l}$, $r_{B,l}$ be the observable number of runs of length l above and below the mean, and m_A, m_B the total number of runs of all lengths above and below the mean, respectively. If r is the total number of runs as defined in 4.1, then

$$r = m_A + m_B + 1 \tag{4.2.1}$$

where

$$m_A = \sum_{l=1}^{\max l} r_{A,l}$$

and

$$m_B = \sum_{l=1}^{\max l} r_{B,l}$$

The theoretical expectations of $r_{A,l}$ and $r_{B,l}$ for any l may be easily computed by the following relation (ref. 8):

$$E(r_{i,l}) = \frac{N - l + 3}{2^{l+2}} \qquad i = A,B \tag{4.2.2}$$

where N is the number of elements in each generated sequence. An arbitrary size $N = 10{,}000$ was chosen for each sequence to be tested for this study; therefore, the expectations of runs above or below the mean follow:

$$\begin{array}{lll}
E(r_{i,l=1}) = 1250.25 & E(r_{i,l=5}) = 78.11 & E(r_{i,l=9}) = 4.88 \\
E(r_{i,l=2}) = 625.06 & E(r_{i,l=6}) = 39.05 & E(r_{i,l=10}) = 2.44 \\
E(r_{i,l=3}) = 312.50 & E(r_{i,l=7}) = 19.52 & E(r_{i,l=11}) = 1.22 \\
E(r_{i,l=4}) = 156.23 & E(r_{i,l=8}) = 9.76 & E(r_{i,l\geq 12}) \cong 1.00
\end{array}$$

$$i = A,B$$

Consider testing the null hypothesis that the deviations between the observed number of runs of any length and their corresponding theoretical expectations are fairly small. By this hypothesis one implies that if the difference for each and every length between observed and expected is small, it is reasonable to believe that an indication of randomness is present for a given sequence of N observations. To test such a hypothesis, it has been the practice of many to use Karl Pearson's chi-square goodness-of-fit test (ref. 16). (For a fuller discussion of this test, see Chapter V.) Notice here that one's interest is not to test the null hypothesis that the observed number of runs of the various lengths has some specified distribution function. Rather, the primary interest is to determine whether the difference between observed and expected is significant to warrant the existence of nonrandomness in a given sequence of observations.

Let $r_{i,l}$ be the observed number of runs of length l, and let the corresponding theoretical numbers of runs that should be in the l^{th} class be $E(r_{i,l})$. Then from the $k = 12$ classes, the test statistic

$$\tau = \sum_{l=1}^{k} \frac{\left[r_{i,l} - E(r_{i,l})\right]^2}{E(r_{i,l})} \qquad i = A,B \tag{4.2.3}$$

is approximately distributed as a chi-square with $k - 1$ degrees of freedom since no parameter estimation is needed. Hence, to test the null hypothesis in question, the value of τ is determined and compared with the upper tail of the chi-square distribution with $k - 1 = 11$ degrees of freedom and a given level of significance. The null hypothesis is rejected whenever τ exceeds this critical region. Rejection would be sufficient to suspect nonrandomness in a given sequence of numbers.

This test was computed for runs above or below the mean for each of the 200 sequences. The results appear in tables (4.20), (4.21), and (4.22). The apparent conclusions are the following:

(a) For the 100 sequences generated by equation (2.1.4), there were six rejections for runs above and seven for runs below of all lengths. This number of rejections is to be expected. Hence, there is an indication that equation (2.1.4) generates sequences that appear to be random based on this test.

(b) For the 50 sequences generated by equation (2.2.1), there were six rejections for runs above and four for runs below of all lengths. Hence, the apparent conclusion is similar to (a).

(c) For the 50 sequences generated by equation (2.2.2), there were 10 rejections for runs above and only 2 for runs below of all lengths. Because of the significant difference between these two numbers, it would seem to suggest that equation (2.2.2) tends to produce exceedingly more runs than expected of values that are greater than the mean. Hence, some nonrandomness appears to exist.

4.3 Serial Correlation

If a given sequence of numbers is truly random, each element should be independent of any other. That is to say, the correlation between an element x_i in the sequence and another element x_{i+p} should be negligible.

Let $x_1, x_2, \ldots, x_N$ be the sequence to be tested for randomness. Let u_j be equal to x_{i+p} for $i = 1, 2, \ldots, N - p$. For this arrangement, the corresponding values of x and u are indicated in table 1.

TABLE 1

X	x_1	x_2	. . .	x_i	. . .	x_{N-p}
U	x_{1+p}	x_{2+p}	. . .	x_{i+p}	. . .	x_N

The correlation coefficient computed by using the configuration of table 1 is called the serial correlation coefficient with lag p. Standard regression and correlation methods are not applicable here because the u_j's no longer constitute a random sample for any fixed set of x's.

A nonparametric method based on serial correlation can be derived if one assumes that all possible permutations of a given sequence of numbers are equally probable. However, the number of permutations become extremely numerous for N at all large; therefore, it is necessary to use an approximation for the distribution of the serial correlation coefficient when N is large.

Referring to table 1, $N - p$ pairs of elements can be formed. Based on these pairs, the serial correlation coefficient may be expressed in the form

$$\rho = \frac{\sum_{i=1}^{N-p} x_i x_{i+p} - (N - p)\bar{x}\bar{u}}{(N - p)s_x s_u} \tag{4.3.1}$$

where

$$\bar{x} = \sum_{i=1}^{N-p} \frac{x_i}{N - p}$$

$$\bar{u} = \sum_{i=1}^{N-p} \frac{x_{i+p}}{N - p}$$

$$s_x^2 = \frac{\sum_{i=1}^{N-p} (x_i - \bar{x})^2}{N - p - 1}$$

$$s_u^2 = \frac{\sum_{i=1}^{N-p} (x_{i+p} - \bar{u})^2}{N - p - 1}$$

For this study, serial correlation coefficients of lags one through fifteen, $p = 1, 2, \ldots, 15$, have been computed. This range is believed to be sufficient to detect any existence of interdependence among the elements of a sequence being tested.

Consider the use of some approximation for the distribution of the serial correlation coefficient. For $N = 10{,}000$, it is doubtful whether the quantities $\bar{x}$, $\bar{u}$, s_x, s_u will be appreciably altered for any lag $p = 1, 2, \ldots, 15$. However, the quantity that will be affected significantly from one lag to another is the sum $\sum_{i=1}^{N-p} x_i x_{i+p}$; therefore, it would be more beneficial to study the distribution of this sum rather than the distribution of ρ itself. Let

$$R = \sum_{i=1}^{N-p} x_i x_{i+p} \tag{4.3.2}$$

Assume that the elements of a generated sequence being tested constitute a random sample from a distribution possessing low order moments - which is the case here - then A. Wald and J. Wolfowitz (ref. 9) have shown that R is a random variable approximately distributed as a normal for large N with mean

$$E(R) = \frac{S_1^2 - S_2}{N - 1} \qquad (4.3.3)$$

and variance

$$\mathrm{var}(R) = \frac{S_2^2 - S_4}{N - 1} + \frac{S_1^4 - 4S_1^2 S_2 + 4S_1 S_3 + S_2^2 - 2S_4}{(N - 1)(N - 2)} - E^2(R) \qquad (4.3.4)$$

where

$$S_k = \sum_{i=1}^{N} x_i^k \qquad (4.3.5)$$

Thus, to test the null hypothesis of zero serial correlation consider the standard normal random variable

$$Z = \frac{R - E(R)}{\sigma_R} \qquad (4.3.6)$$

Calculate and compare Z to the left as well as the right tail of the standard normal - with a probability 0.05 for Type I error - because both large positive or large negative serial correlations are of interest. Reject the hypothesis if Z exceeds these limits and conclude that some element dependence is present.

The serial correlation test is sensitive to periodicities caused by the dependence between elements p spaces apart and should offset any deficiency in the runs test. As before, all sequences were subjected to this test with the results appearing in tables (4.30), (4.31), and (4.32). The results indicate the following:

(a) For the 100 sequences generated by equation (2.1.4), no one lag was rejected significantly more than any other. The most rejections noted for any one lag of the fifteen considered were ten for lag p equal to two. The overall average number of rejections for any one lag was six percent, and this, of course, is within the framework of the statistical test. Hence, it is apparent that no periodicities caused by the dependence between elements seem to exist.

(b) For the 50 sequences generated by equation (2.2.1), the most rejections noted for any one lag were four with the overall number of rejections for any one lag being approximately equal to five percent. Hence, for the normal random number generator

$$X_1 = (-2 \ln U_1)^{1/2} \sin 2\pi U_2$$

no element dependence appears to exist.

(c) Quite a different result was noted for the sequences generated by equation (2.2.2). An overwhelming number of 41 rejections occurred for lag p equal to one. This fact indicates strongly the existence of element dependence one space apart. Hence, it would be reasonable to conclude that the normal random number generator

$$X_2 = (-2 \ln U_1)^{1/2} \cos 2\pi U_2$$

contains serious types of nonrandomness caused by element dependence.

4.4 Numerical Results

All tests described in this chapter were performed on each and every one of the two-hundred (200) sequences generated for this study. All necessary calculations were made on the IBM-7094 computer with the results of the tests having been rounded for presentation. Rejection is indicated by an asterisk. The critical values for the appropriate random variables may be found under each table. Each table is identified with respect to its contents.

TABLE 4.10

RESULTS OF THE TEST ON TOTAL RUNS OF 100 SEQUENCES OF SIZE 10,000 EACH

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4)

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
1	5052	5000.82	50.00	1.024
2	5019	5001.00	50.00	0.360
3	4960	5000.44	49.99	-0.809
4	5040	5000.92	50.00	0.782
5	5044	5000.97	50.00	0.861
6	5010	4995.01	49.94	0.300
7	5051	5000.81	50.00	1.004
8	5008	5000.74	49.99	0.145
9	4997	5000.58	49.99	-0.072
10	5013	5000.02	49.99	0.260
11	5002	5000.18	49.99	0.036
12	5004	5000.96	50.00	0.061
13	5110	5000.90	50.00	2.182*
14	4988	5000.90	50.00	-0.258
15	5024	5000.99	50.00	0.460
16	5010	5001.00	50.00	0.180
17	5043	5000.68	49.99	0.846

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.10

RESULTS OF THE TEST ON TOTAL RUNS OF 100 SEQUENCES OF SIZE 10,000 EACH

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
18	4949	5000.99	50.00	-1.040
19	4986	4999.34	49.98	-0.267
20	5000	5001.00	50.00	-0.020
21	4979	5001.00	50.00	-0.440
22	5022	5000.80	50.00	0.424
23	4998	5000.63	49.99	-0.053
24	5038	5000.30	49.99	0.754
25	4960	5000.44	49.99	-0.809
26	5063	5000.70	49.99	1.246
27	4974	5000.30	49.99	-0.526
28	5024	5000.48	49.99	0.470
29	5060	5001.00	50.00	1.180
30	5048	5000.92	50.00	0.942
31	4937	5001.00	50.00	-1.280
32	4998	5001.00	50.00	-0.060
33	5023	4998.88	49.98	0.483
34	5056	5000.82	50.00	1.104

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.10

RESULTS OF THE TEST ON TOTAL RUNS OF 100 SEQUENCES OF SIZE 10,000 EACH

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
35	4979	5000.99	50.00	-0.440
36	5029	5000.50	49.99	0.570
37	5019	5000.13	49.99	0.378
38	5064	5000.99	50.00	1.260
39	4978	4995.08	49.94	-0.342
40	4939	5000.96	50.00	-1.239
41	5026	5000.08	49.99	0.519
42	5021	5000.97	50.00	0.401
43	4961	5001.00	50.00	-0.800
44	5012	5001.00	50.00	0.220
45	5036	5000.77	50.00	0.705
46	4976	5000.93	50.00	-0.499
47	5017	5000.65	49.99	0.327
48	5027	4999.84	49.99	0.543
49	5001	4995.15	49.94	0.117
50	4999	5000.35	49.99	-0.027
51	5039	5000.65	49.99	0.767

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.10

RESULTS OF THE TEST ON TOTAL RUNS OF 100 SEQUENCES OF SIZE 10,000 EACH

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
52	5064	5000.78	50.00	1.264
53	4966	5001.00	50.00	-0.700
54	4956	5000.99	50.00	-0.900
55	4940	5000.73	49.99	-1.215
56	5040	5000.30	49.99	0.794
57	5005	5001.00	50.00	0.080
58	4951	5000.82	50.00	-0.996
59	4933	5000.99	50.00	-1.360
60	5041	5000.94	50.00	0.801
61	5173	5000.85	50.00	3.443*
62	5160	5000.94	50.00	3.181*
63	5010	5001.00	50.00	0.180
64	5019	5000.97	50.00	0.361
65	5059	4999.72	49.98	1.186
66	4986	5001.00	50.00	-0.300
67	5037	5000.93	50.00	0.721
68	5005	5000.98	50.00	0.080

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.10

RESULTS OF THE TEST ON TOTAL RUNS OF 100 SEQUENCES OF SIZE 10,000 EACH GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
69	5002	5000.74	49.99	0.025
70	5033	5000.28	49.99	0.655
71	4975	5000.88	50.00	-0.518
72	5045	5000.97	50.00	0.881
73	4980	5001.00	50.00	-0.420
74	5069	4999.81	49.99	1.384
75	4980	5000.82	50.00	-0.416
76	5026	5000.50	49.99	0.510
77	4953	5000.74	49.99	-0.955
78	4992	5000.90	50.00	-0.178
79	5022	5000.66	49.99	0.427
80	5088	5000.94	50.00	1.741
81	4996	5000.97	50.00	-0.099
82	4980	5000.99	50.00	-0.420
83	5000	5000.26	49.99	-0.005
84	4974	5000.99	50.00	-0.540
85	5027	5000.98	50.00	0.521

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.10

RESULTS OF THE TEST ON TOTAL RUNS OF 100 SEQUENCES OF SIZE 10,000 EACH GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Concluded

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
86	5067	5000.73	49.99	1.326
87	5116	5000.77	50.00	2.305*
88	5074	5000.84	50.00	1.463
89	5003	4999.19	49.98	0.076
90	5049	4999.72	49.98	0.986
91	4993	5000.70	49.99	-0.154
92	5008	5000.84	50.00	0.143
93	4992	5000.92	50.00	-0.178
94	4950	5001.00	50.00	-1.020
95	4987	5000.98	50.00	-0.280
96	5023	5001.00	50.00	0.440
97	5006	5000.02	49.99	0.120
98	5029	5000.18	49.99	0.577
99	4978	4999.38	49.98	-0.428
100	5024	5000.86	50.00	0.463

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.11

RESULTS OF THE TEST ON TOTAL RUNS OF 50 SEQUENCES OF SIZE 10,000

EACH GENERATED BY EQUATION (2.2.1)

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
1	4956	5000.99	50.00	-0.900
2	4966	5001.00	50.00	-0.700
3	5064	5000.78	50.00	1.264
4	5039	5000.65	49.99	0.767
5	4987	5000.99	50.00	-0.280
6	5044	5000.81	50.00	0.864
7	5005	5000.98	50.00	0.080
8	4987	5000.98	50.00	-0.280
9	4981	5001.00	50.00	-0.400
10	5058	5000.78	50.00	1.144
11	5031	5000.44	49.99	0.611
12	5023	5001.00	50.00	0.440
13	4983	5000.96	50.00	-0.359
14	5056	5000.84	50.00	1.103
15	5006	5000.02	49.99	0.120
16	5000	5001.00	50.00	-0.020
17	5029	5000.18	49.00	0.577

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.11

RESULTS OF THE TEST ON TOTAL RUNS OF 50 SEQUENCES OF SIZE 10,000

EACH GENERATED BY EQUATION (2.2.1) - Continued

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
18	4994	5000.94	50.00	-0.139
19	4976	5000.61	49.99	-0.492
20	4995	5000.80	50.00	-0.116
21	4978	4999.38	49.98	-0.428
22	5072	5000.99	50.00	1.420
23	5024	5000.86	50.00	0.463
24	4962	5000.35	49.99	-0.767
25	4940	5000.73	49.99	-1.215
26	5040	5000.30	49.99	0.794
27	5005	5001.00	50.00	0.080
28	4951	5000.82	50.00	-0.996
29	4933	5000.99	50.00	-1.360
30	5041	5000.94	50.00	0.801
31	5173	5000.85	50.00	3.443*
32	5160	5000.94	50.00	3.181*
33	5010	5001.00	50.00	0.180
34	5019	5000.97	50.00	0.361

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.11

RESULTS OF THE TEST ON TOTAL RUNS OF 50 SEQUENCES OF SIZE 10,000

EACH GENERATED BY EQUATION (2.2.1) - Concluded

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
35	5059	4999.72	49.98	1.186
36	4986	5001.00	50.00	-0.300
37	5037	5000.93	50.00	0.721
38	5005	5000.98	50.00	0.080
39	5002	5000.74	49.99	0.025
40	5033	5000.28	49.99	0.655
41	4975	5000.88	50.00	-0.518
42	5045	5000.97	50.00	0.881
43	4980	5001.00	50.00	-0.420
44	5069	4999.81	49.99	1.384
45	5116	5000.77	50.00	2.305*
46	5000	5000.26	49.99	-0.005
47	4974	5000.99	50.00	-0.540
48	4996	5000.97	50.00	-0.099
49	5022	5000.66	49.99	0.427
50	4950	5001.00	50.00	-1.020

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.12

RESULTS OF THE TEST ON TOTAL RUNS OF 50 SEQUENCES OF SIZE 10,000 EACH GENERATED BY EQUATION (2.2.2)

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
1	4993	5000.98	50.00	-0.160
2	5018	5000.86	50.00	0.343
3	4997	5000.99	50.00	-0.080
4	5004	5000.88	50.00	0.062
5	4949	4999.47	49.98	-1.010
6	4968	5000.90	50.00	-0.658
7	4957	5000.71	49.99	-0.874
8	4947	4999.27	49.98	-1.046
9	4995	5000.70	49.99	-0.114
10	4965	5000.92	50.00	-0.718
11	4993	5000.93	50.00	-0.159
12	4959	5000.77	50.00	-0.835
13	4965	5000.93	50.00	-0.719
14	5065	5000.92	50.00	1.282
15	5020	5000.88	50.00	0.383
16	4986	5000.86	50.00	-0.297
17	5073	5000.97	50.00	1.441

Critical values for the 0.05 probability level are ± 1.96.

TABLE 4.12

RESULTS OF THE TEST ON TOTAL RUNS OF 50 SEQUENCES OF SIZE 10,000 EACH GENERATED BY EQUATION (2.2.2) - Continued

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
18	5023	5000.99	50.00	0.440
19	5036	4998.88	49.98	0.743
20	4962	5000.93	50.00	-0.779
21	4943	5000.33	49.99	-1.147
22	5062	5000.65	49.99	1.227
23	4996	5000.16	49.99	-0.083
24	5025	5000.94	50.00	0.481
25	5024	5000.95	50.00	0.461
26	4996	5001.00	50.00	-0.100
27	5043	5000.98	50.00	0.841
28	5003	5000.98	50.00	0.040
29	5008	5000.89	50.00	0.142
30	5065	5000.82	50.00	1.284
31	5011	5000.99	50.00	0.200
32	4959	5000.65	49.99	-0.833
33	5035	5000.16	49.99	0.697
34	5014	5000.88	50.00	0.262

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.12

RESULTS OF THE TEST ON TOTAL RUNS OF 50 SEQUENCES OF SIZE 10,000

EACH GENERATED BY EQUATION (2.2.2) - Concluded

Sequence no.	r (Obs. no. of runs)	μ_r (Mean of runs)	σ_r (S.D. of runs)	$Z = \frac{r - \mu_r}{\sigma_r}$
35	5019	5000.33	49.99	0.374
36	4987	5001.00	50.00	-0.280
37	5037	5000.60	49.99	0.728
38	4960	5000.99	50.00	-0.820
39	5048	5000.10	49.99	0.958
40	5019	5000.68	49.99	0.366
41	4913	5000.75	50.00	-1.755
42	5065	5001.00	50.00	1.280
43	5079	5000.50	49.99	1.570
44	4998	5000.77	50.00	-0.055
45	5011	5000.63	49.99	0.207
46	4976	5000.80	50.00	-0.496
47	5067	5000.74	49.99	1.325
48	5043	4999.59	49.98	0.869
49	5042	4999.45	49.98	0.851
50	4928	5000.97	50.00	-1.460

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.20

RESULTS OF THE TEST ON RUNS ABOVE AND BELOW THE MEAN FOR 100 SEQUENCES GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4)

Sequence No.

		1	2	3	4	5	6	7	8	9	10
χ^2	Above	8.759	3.834	14.523	11.293	10.324	11.520	11.609	3.418	3.602	13.260
	Below	9.624	2.848	16.348	8.201	15.352	12.936	9.690	5.286	12.887	10.475

Sequence No.

		11	12	13	14	15	16	17	18	19	20
χ^2	Above	19.468	8.132	18.543	15.896	5.176	8.184	9.313	9.043	14.349	3.068
	Below	6.733	13.497	14.945	25.563*	4.865	11.820	16.547	16.777	16.048	20.651*

Sequence No.

		21	22	23	24	25	26	27	28	29	30
χ^2	Above	8.741	13.265	5.807	11.360	14.523	12.065	14.598	8.482	6.496	8.013
	Below	16.526	5.535	16.173	5.750	16.348	27.605*	7.685	13.336	11.609	14.627

Sequence No.

		31	32	33	34	35	36	37	38	39	40
χ^2	Above	6.496	2.779	10.998	16.552	8.426	13.136	4.692	10.434	11.971	11.627
	Below	10.856	21.518*	13.301	15.471	15.399	11.096	16.759	7.935	15.011	7.566

Sequence No.

		41	42	43	44	45	46	47	48	49	50
χ^2	Above	10.975	5.013	13.400	22.626*	11.410	15.045	11.936	11.735	12.167	20.316*
	Below	10.412	5.514	6.595	8.235	9.262	5.758	11.251	8.378	13.141	6.607

Critical value of χ^2 for 0.05 probability level with 11 degrees of freedom is 19.7.

TABLE 4.20

RESULTS OF THE TEST ON RUNS ABOVE AND BELOW THE MEAN FOR 100 SEQUENCES GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Concluded

Sequence No.

		51	52	53	54	55	56	57	58	59	60
χ^2	Above	12.310	3.272	11.540	8.013	11.460	11.774	9.622	9.091	4.194	8.896
	Below	10.462	8.613	7.743	9.570	15.433	5.793	23.506*	16.863	9.741	15.155

Sequence No.

		61	62	63	64	65	66	67	68	69	70
χ^2	Above	17.840	15.037	3.657	5.888	4.575	6.057	6.434	3.392	2.934	11.297
	Below	14.194	17.395	3.838	5.447	10.925	23.540*	5.006	1.949	14.066	5.967

Sequence No.

		71	72	73	74	75	76	77	78	79	80
χ^2	Above	7.905	6.459	11.741	6.903	5.354	4.940	18.025	22.229*	12.183	16.436
	Below	11.008	5.146	26.162*	12.107	7.715	4.606	9.706	11.550	10.952	16.683

Sequence No.

		81	82	83	84	85	86	87	88	89	90
χ^2	Above	25.128*	9.802	21.194*	12.617	7.890	9.455	18.950	3.594	23.970*	6.911
	Below	12.201	15.287	6.130	14.277	13.128	9.943	18.661	9.792	6.744	7.598

Sequence No.

		91	92	93	94	95	96	97	98	99	100
χ^2	Above	3.868	5.786	6.337	4.411	9.372	5.401	6.730	6.255	13.968	13.308
	Below	14.595	5.716	5.618	7.922	12.565	6.255	16.441	5.904	16.980	6.023

Critical value of χ^2 for 0.05 probability level with 11 degrees of freedom is 19.7.

TABLE 4.21

RESULTS OF THE TEST ON RUNS ABOVE AND BELOW THE MEAN FOR 50 SEQUENCES GENERATED BY EQUATION (2.2.1)

Sequence No.

		1	2	3	4	5	6	7	8	9	10
χ^2	Above	9.570	7.743	8.613	10.462	21.784*	15.290	27.020*	12.563	20.684*	15.913
	Below	8.013	11.540	3.272	12.310	8.086	8.878	14.969	9.272	4.041	16.724

Sequence No.

		11	12	13	14	15	16	17	18	19	20
χ^2	Above	10.994	6.235	16.994	15.135	16.441	21.441*	5.904	12.479	18.423	9.975
	Below	13.181	5.401	7.616	17.860	6.730	2.948	6.255	21.615*	7.076	33.566*

Sequence No.

		21	22	23	24	25	26	27	28	29	30
χ^2	Above	16.980	13.028	6.023	16.458	15.433	5.793	23.506*	16.863	9.741	15.155
	Below	13.968	6.119	13.308	14.780	11.460	11.774	9.622	9.091	4.194	8.896

Sequence No.

		31	32	33	34	35	36	37	38	39	40
χ^2	Above	14.194	17.395	3.838	5.447	10.925	12.540	5.006	1.949	14.066	5.967
	Below	17.840	15.037	3.657	5.888	4.575	6.057	6.434	3.392	2.934	11.297

Sequence No.

		41	42	43	44	45	46	47	48	49	50
χ^2	Above	11.008	5.146	26.162*	12.107	18.661	6.130	14.277	12.201	10.952	7.922
	Below	7.905	6.459	11.741	6.903	18.950	21.194*	12.617	25.128*	12.183	4.411

Critical value of χ^2 for 0.05 probability level with 11 degrees of freedom is 19.7.

TABLE 4.22

RESULTS OF THE TEST ON RUNS ABOVE AND BELOW THE MEAN OF 50 SEQUENCES GENERATED BY EQUATION (2.2.2)

Sequence No.

		1	2	3	4	5	6	7	8	9	10
χ^2	Above	16.679	6.353	10.830	13.916	25.485*	7.880	14.154	25.226*	17.955	21.598*
	Below	8.198	13.743	9.500	5.595	13.349	6.670	4.178	13.512	13.677	15.616

Sequence No.

		11	12	13	14	15	16	17	18	19	20
χ^2	Above	14.856	19.079	23.353*	7.958	14.529	22.435*	7.884	17.123	15.014	8.536
	Below	8.036	7.415	8.686	7.664	5.163	12.298	9.256	7.893	12.584	9.106

Sequence No.

		21	22	23	24	25	26	27	28	29	30
χ^2	Above	23.719*	9.205	27.595*	15.081	5.730	22.715*	9.646	17.790	9.024	6.641
	Below	16.275	6.888	12.481	4.808	15.036	7.548	7.662	24.815*	5.432	9.007

Sequence No.

		31	32	33	34	35	36	37	38	39	40
χ^2	Above	18.070	9.130	14.137	17.578	10.023	7.062	13.512	8.050	17.063	12.316
	Below	21.055*	6.052	8.569	13.031	4.097	11.451	7.425	9.160	8.764	6.270

Sequence No.

		41	42	43	44	45	46	47	48	49	50
χ^2	Above	25.316*	9.330	11.163	17.682	14.842	10.974	10.204	17.199	17.716	23.454*
	Below	17.151	11.145	10.337	10.851	10.918	7.487	7.710	16.276	16.600	6.087

Critical value of χ^2 for 0.05 probability level with 11 degrees of freedom is 19.7.

TABLE 4.30

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 100 SEQUENCES

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4)

Seq. No. / Lag p	1	2	3	4	5	6	7	8	9	10
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-1.377	-0.969	1.237	-1.045	-1.082	-1.015	-0.835	-0.769	-0.007	-1.276
2	-0.282	0.123	-2.698*	-1.357	1.022	-1.307	-2.443*	0.286	1.338	-0.910
3	-0.321	-1.699	1.092	0.342	0.091	0.039	1.438	-1.300	-1.151	1.403
4	-0.300	0.733	-1.861	-0.083	-0.973	-0.706	0.742	0.592	0.438	-1.136
5	-0.531	0.042	-0.103	-1.264	-0.541	-0.120	-0.776	-0.057	0.235	0.186
6	-1.101	0.334	-1.690	-0.611	-1.701	2.115*	0.210	-0.390	0.263	1.475
7	-1.604	1.192	0.453	-0.504	1.341	-0.428	0.126	1.376	-0.956	-0.316
8	0.412	-1.886	0.171	0.191	-2.586*	-1.889	1.123	-1.453	-0.488	-1.198
9	-1.979*	-1.811	-0.004	-0.755	1.573	1.209	-1.676	-1.999*	-0.767	-0.919
10	1.028	-0.262	-2.183*	-0.379	-0.760	-0.564	-1.009	-0.142	-0.847	-0.250
11	-1.131	0.318	0.521	-0.713	1.384	-1.397	0.157	-0.002	-0.585	-0.449
12	-0.140	0.051	0.285	0.600	0.265	-1.473	-0.556	0.282	-0.738	2.480*
13	0.595	-0.242	0.269	-1.644	-1.241	0.974	-0.119	-0.053	-0.332	0.862
14	-0.120	-0.064	0.347	1.463	1.918	-1.422	0.505	0.447	-0.470	1.511
15	-1.304	0.137	0.437	-1.489	-1.870	0.208	-0.887	-0.031	1.105	-1.739

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.30

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 100 SEQUENCES

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Seq. No. / Lag p	11	12	13	14	15	16	17	18	19	20
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-0.450	-1.289	-2.314*	0.027	-0.415	-1.437	-1.203	0.466	0.595	-1.253
2	-0.178	-0.561	-2.275*	-0.919	-1.351	-0.864	0.628	-1.801	-2.941*	-0.560
3	-0.888	0.297	-0.507	-0.920	-0.780	0.159	0.392	0.213	1.296	-1.331
4	0.218	-1.046	0.138	0.606	-0.464	-1.256	-1.213	-0.515	-1.195	-0.735
5	0.277	0.866	-0.256	-1.384	-0.325	0.733	-0.771	-0.856	0.295	0.065
6	0.262	0.819	-0.950	-1.420	0.700	0.976	-1.865	0.924	-1.550	-0.228
7	-1.136	0.664	-0.707	-0.246	0.605	0.044	1.124	-0.983	0.310	2.063*
8	0.865	0.264	-1.421	-0.583	-1.591	0.072	-2.957*	1.646	-0.038	0.250
9	-1.920	0.149	-0.016	-1.813	0.019	0.366	1.647	0.895	-0.664	-1.438
10	-1.853	-0.260	-1.628	1.008	1.555	-0.405	-1.069	-1.585	-2.429*	-1.518
11	-1.288	-0.328	1.159	-0.194	-1.707	-0.207	1.271	-0.059	0.593	0.808
12	1.119	1.183	-1.935	0.474	-0.496	1.376	0.087	0.434	0.673	-0.471
13	1.005	-1.940	0.760	-0.376	-0.406	-1.861	-1.311	-1.803	0.318	-0.553
14	-0.951	0.189	-0.913	-2.303*	-1.445	-0.054	1.904	-1.952	-0.426	0.226
15	-0.393	-0.360	-1.642	-0.128	1.449	-0.701	-1.768	1.404	0.327	1.172

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.30

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 100 SEQUENCES

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Seq. No. / Lag p	21	22	23	24	25	26	27	28	29	30
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-0.759	-0.784	-0.252	-1.864	1.237	-2.285*	-0.990	-0.764	-2.069*	-1.206
2	-1.963*	-1.522	0.912	-0.344	-2.698*	-2.355*	-2.737*	-0.343	-0.438	1.021
3	-0.891	-0.290	-0.978	1.325	1.092	-0.057	-0.497	-0.225	0.323	0.001
4	0.695	-0.188	-1.598	0.406	-1.861	1.939	0.539	0.023	-1.056	-1.228
5	-0.094	-0.613	-0.159	0.163	-0.103	-0.356	-0.829	-0.774	0.807	-0.593
6	-2.122*	-0.641	0.097	1.259	-1.690	-0.913	-1.707	-0.291	0.389	-1.647
7	0.596	-0.010	-0.524	-1.115	0.453	-0.573	-0.540	-0.573	-0.048	1.150
8	0.067	0.575	-1.649	-0.305	0.171	-0.784	-0.052	-1.028	-0.045	-2.751*
9	-0.723	-0.946	-0.646	-0.879	-0.004	-0.281	-0.455	0.006	0.184	1.656
10	0.297	-0.598	-0.540	0.214	-2.183*	-1.070	-0.083	-2.166*	-0.475	-1.007
11	-0.670	-1.324	-1.093	-0.406	0.521	0.926	-1.182	0.328	-0.413	1.532
12	-0.132	-0.022	-2.032*	1.912	0.285	-1.077	-0.382	-0.753	0.926	0.228
13	0.605	-1.724	-1.563	0.622	0.269	0.764	1.073	-2.077*	-1.663	-1.213
14	-1.526	0.712	-0.840	1.966*	0.347	-0.980	-0.768	-0.900	-0.202	1.934
15	-0.143	-1.264	1.269	-0.062	0.437	0.410	-0.760	0.383	-0.774	-1.689

Critical values for the 0.05 probability level are +1.96.

TABLE 4.30

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 100 SEQUENCES

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Seq. No. / Lag p	31	32	33	34	35	36	37	38	39	40
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	0.471	-1.348	-1.530	-2.288*	0.242	-0.561	-0.392	-1.535	-0.349	0.503
2	-1.033	-0.666	-0.804	1.510	-1.855	0.291	-0.894	0.374	-1.156	-0.986
3	0.073	-1.261	0.882	2.105*	0.170	-0.815	-0.668	1.162	-0.366	-0.327
4	1.144	-0.613	-1.459	-1.308	-0.349	-2.428*	-0.149	-0.651	-0.279	-0.882
5	-0.119	-0.311	0.046	-0.631	-1.040	-0.092	-0.487	-0.149	-0.079	-0.108
6	0.672	-0.225	1.325	0.555	1.174	0.512	-0.103	-0.720	2.327*	-0.004
7	-1.190	2.035*	-0.353	-1.610	-0.921	1.212	1.441	-0.461	-0.458	-1.917
8	-1.255	0.230	-1.378	-1.613	0.902	-1.786	-0.917	-1.507	-2.262*	0.320
9	0.412	-1.601	-1.116	1.383	0.636	0.010	-0.857	3.111*	0.690	0.210
10	-0.340	-1.476	-0.621	-1.087	-1.994*	-1.049	0.543	-0.050	-1.025	0.301
11	-2.850*	0.887	-0.364	0.179	-0.358	0.867	0.385	-1.384	-1.508	0.115
12	0.193	-0.648	2.262*	-0.120	0.706	-1.145	1.195	-0.116	-1.288	1.201
13	-0.253	-0.433	1.347	-0.511	-1.320	-0.445	-1.060	-0.624	0.579	0.029
14	-0.493	0.335	1.861	-0.472	-2.241*	-2.232*	-0.361	-0.942	-1.121	-0.928
15	1.310	1.272	-2.051*	0.317	1.104	-1.639	-0.870	-0.541	-0.467	-1.080

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.30

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 100 SEQUENCES

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Seq. No. / Lag p	41	42	43	44	45	46	47	48	49	50
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-1.448	-0.300	0.541	-1.487	-1.058	0.254	-1.219	-1.682	-0.717	-0.040
2	-0.922	-1.484	-0.460	1.462	-1.381	-0.279	-1.254	-0.817	-1.565	-0.408
3	1.436	-0.715	-1.233	0.157	0.216	-1.232	-0.043	1.404	-0.302	-0.918
4	-1.064	-0.221	-0.086	-1.166	-0.034	-0.491	-0.164	-0.276	-0.761	0.439
5	-0.021	-0.274	-1.412	-0.297	-1.530	-1.431	-0.975	-0.381	-0.316	0.339
6	1.293	0.873	-1.338	0.986	-0.511	-0.820	-0.570	0.838	2.161*	0.723
7	-0.268	0.684	0.571	-2.263*	-0.704	0.951	-0.640	-1.105	-0.371	-1.151
8	-0.931	-1.614	-2.053*	-0.437	0.040	-2.081*	0.189	-0.698	-2.323*	0.323
9	-0.806	0.049	-0.434	1.443	-0.673	-0.528	-0.862	-1.199	0.652	-2.216*
10	-0.290	1.823	-1.368	-0.257	-0.456	-1.302	-0.693	0.006	-1.011	-1.649
11	-0.282	-1.911	0.954	0.014	-1.123	0.830	-1.335	0.003	-1.413	-0.909
12	2.625*	-0.634	-0.348	0.318	0.491	0.067	0.333	2.051*	-1.306	0.552
13	0.967	-0.562	0.453	-0.363	-1.842	0.398	-1.458	0.583	0.778	0.534
14	1.609	-1.523	-2.285*	-1.115	1.251	-1.973*	1.027	1.394	-1.225	-0.994
15	-1.710	1.466	-1.343	0.379	-1.829	-1.322	-2.098*	-0.856	-0.267	-0.501

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.30

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 100 SEQUENCES

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Seq. No. / Lag p	51	52	53	54	55	56	57	58	59	60
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-0.780	-0.885	0.161	-0.546	0.599	-1.836	-1.216	0.431	0.585	-1.152
2	0.301	-1.444	-0.791	-0.331	-0.265	-0.408	-0.669	-0.573	-1.740	0.742
3	-0.515	0.345	0.459	0.048	1.001	1.513	-1.190	0.702	0.506	0.694
4	-2.222*	1.016	-1.145	-2.358*	1.530	0.513	-1.732	1.902	-1.155	-1.146
5	0.085	-1.145	0.025	1.412	-0.071	0.190	-1.021	-0.162	-1.208	-0.619
6	0.651	0.074	-0.190	-0.689	-0.205	1.367	0.146	-0.079	-0.149	-1.932
7	1.194	0.131	-1.999*	-2.306*	-1.460	-1.110	1.188	-1.536	-0.163	1.220
8	-1.500	1.547	0.761	-0.184	-0.974	-0.236	-0.863	-0.871	-1.941	-3.121*
9	-0.227	-1.132	0.262	0.829	0.630	-0.558	-1.578	0.854	0.442	2.056*
10	-1.136	-0.899	0.057	0.533	0.062	0.332	-1.527	-0.066	-1.078	-1.283
11	0.800	-0.178	0.380	0.104	-2.908*	-0.339	0.447	-2.554*	-0.199	1.297
12	-0.947	-1.207	1.170	0.082	0.154	2.201*	-1.090	0.399	-0.943	-0.126
13	-0.293	-0.435	-0.033	-0.408	-0.260	0.839	-1.011	0.274	0.521	-1.045
14	-2.192*	0.190	-0.755	-0.650	0.075	1.788	0.236	0.367	0.568	1.187
15	-1.714	-1.197	-1.137	0.509	1.207	0.054	0.180	1.188	0.108	-1.814

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.30

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 100 SEQUENCES

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Seq. No. / Lag p	61	62	63	64	65	66	67	68	69	70
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-2.833*	-2.716*	-0.896	-0.290	-1.228	0.422	-0.730	-0.430	-0.076	-1.765
2	0.214	0.772	0.179	-1.449	-1.270	0.008	-1.223	-1.504	1.598	-0.376
3	-0.097	-0.511	-1.581	-0.732	0.083	0.172	-0.893	1.005	-1.188	1.248
4	-0.351	-0.361	0.495	-0.262	0.666	-0.056	-0.604	-1.382	0.262	0.509
5	0.858	0.735	-0.151	-0.248	-0.952	1.574	-0.236	-1.690	0.248	-0.087
6	0.169	0.999	0.246	0.827	-0.937	-2.317*	0.660	-0.256	0.514	0.863
7	-0.409	-1.060	1.585	0.595	-0.093	0.744	0.657	0.684	-0.713	-1.216
8	-1.302	-1.400	-1.433	-1.595	-0.518	-0.989	-1.692	-0.711	-0.542	-0.511
9	-1.220	-0.517	-1.749	0.071	-0.968	0.302	0.044	0.150	-0.598	-0.871
10	1.456	1.189	-0.527	1.651	-1.123	-1.510	1.409	-1.353	-0.859	0.327
11	-0.931	-0.756	0.019	-2.004*	-1.489	-1.778	-1.786	-1.048	-0.688	-0.409
12	-0.508	-0.857	0.002	-0.550	1.214	0.243	-0.522	-1.260	-0.801	2.143*
13	1.224	1.334	-0.282	-0.425	1.519	-0.314	-0.114	0.181	-0.255	0.834
14	-1.023	-1.302	0.003	-1.631	-1.418	-0.284	-1.512	0.806	-0.340	1.795
15	0.843	0.493	-0.126	1.511	1.480	0.594	1.567	0.150	1.197	-0.396

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.30

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 100 SEQUENCES

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Seq. No. / Lag p	71	72	73	74	75	76	77	78	79	80
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	0.309	-0.674	0.279	-2.240*	0.013	-1.453	1.206	-0.842	-1.177	-1.236
2	-1.077	-0.957	-1.587	-1.932	-0.316	0.614	0.242	-0.206	-1.238	-1.343
3	0.324	-0.854	-0.873	0.006	0.586	-1.497	-0.333	-1.312	0.132	-0.492
4	0.793	-0.098	0.014	0.704	-1.229	0.449	-0.074	-0.882	-0.086	0.592
5	-0.490	0.044	-1.068	-1.127	0.247	0.384	-1.097	0.197	-0.939	-0.361
6	0.378	1.023	-1.308	-1.115	0.096	-0.232	-0.662	1.737	-0.461	-0.737
7	-1.690	0.436	-0.091	-0.494	-2.139*	1.617	0.923	0.488	-0.667	-0.283
8	-1.489	-1.531	-1.127	-0.945	0.529	-1.613	-1.218	0.719	0.106	-0.968
9	1.240	0.266	-1.264	-1.256	0.762	-2.293*	-1.254	-0.932	-0.796	-0.085
10	-0.135	1.586	0.626	-0.892	-0.006	-0.133	-1.277	0.846	-0.591	-0.874
11	-3.070*	-1.652	-0.587	-1.327	0.702	-0.211	0.690	0.077	-1.188	1.404
12	0.669	-0.412	0.051	1.174	1.202	0.227	0.299	-0.360	0.479	-1.742
13	-0.238	0.061	-0.569	1.382	0.084	0.312	-0.081	0.935	-1.399	0.646
14	-0.518	-1.081	-2.410*	-1.170	-1.112	0.177	-1.453	0.040	1.069	-0.518
15	1.971*	2.077*	0.877	1.321	-1.294	-0.551	-0.840	-0.713	-2.235*	-1.001

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.30

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 100 SEQUENCES

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Continued

Lag p \ Seq. No.	81	82	83	84	85	86	87	88	89	90
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-0.858	-0.704	-0.031	0.168	-1.591	-1.236	-1.816	-1.311	-0.020	-2.277*
2	0.123	-2.269*	-0.390	-1.229	-0.494	-1.927	0.313	-1.661	-0.792	-1.008
3	-1.439	-0.734	-0.876	0.478	0.260	1.429	0.578	0.424	-0.603	-0.125
4	-0.585	0.612	0.378	0.137	-0.911	1.063	0.456	1.281	0.715	0.586
5	0.065	-0.135	0.333	-0.853	0.772	-1.989*	-1.711	-0.670	0.399	-1.214
6	1.889	-2.075*	0.745	1.191	0.720	0.876	-0.329	0.341	0.800	-1.485
7	0.474	0.426	-1.138	-0.970	-0.106	0.910	-1.014	0.165	-2.505*	-0.751
8	0.676	-0.064	0.322	0.514	0.022	1.339	1.067	1.309	0.080	-0.508
9	0.252	-0.891	-2.182*	0.450	0.257	-1.938	0.546	-1.027	-1.399	-1.078
10	1.210	0.279	-1.833	-1.652	-0.057	-0.905	-0.817	-0.603	-1.648	-1.135
11	0.142	-0.443	-1.014	-0.098	-0.121	0.303	1.119	0.074	-1.156	-1.205
12	-0.684	-0.085	0.438	0.469	1.107	-0.650	1.965*	-0.511	0.576	0.882
13	0.322	0.450	0.442	-1.278	-1.966*	0.208	-0.502	-0.193	-0.098	1.011
14	0.534	-1.388	-1.162	-1.829	0.133	0.389	0.699	-0.020	-1.010	-1.693
15	-0.626	-0.050	-0.727	1.668	-0.262	-0.652	-0.358	-0.790	0.094	1.402

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.30

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 100 SEQUENCES

GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Concluded

Seq. No. / Lag p	91	92	93	94	95	96	97	98	99	100
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	0.069	0.013	-0.308	0.391	-0.988	-0.237	-0.365	-1.343	0.445	-0.779
2	0.685	-0.733	-0.096	-1.710	-2.482*	-0.877	-0.684	0.403	-1.889	-1.678
3	-0.820	-1.175	-0.280	0.128	-0.324	-0.804	-0.672	-1.457	1.329	-0.137
4	-1.916	-0.501	-2.121*	-1.115	0.876	-0.465	-0.333	0.113	-1.562	-0.099
5	-0.180	0.264	0.234	-1.169	-1.000	0.315	-0.447	0.938	0.107	-0.597
6	-0.091	1.491	-1.215	-0.095	-1.811	1.060	-0.026	0.117	-1.584	-0.615
7	-0.584	0.095	-2.268*	-0.166	0.378	0.502	1.297	1.259	0.212	0.037
8	-1.526	-1.734	1.139	-1.915	-0.089	-1.698	-1.200	-1.683	0.119	0.653
9	-0.764	0.018	0.530	0.333	-1.328	0.167	-0.664	-2.613*	-0.934	-0.829
10	-0.318	1.778	-0.513	-0.917	0.051	1.825	0.669	-0.434	-2.401*	-0.307
11	-1.095	-1.788	1.748	-0.162	-0.943	-1.637	-0.088	0.150	0.740	-1.185
12	-1.978*	-0.142	1.258	-0.822	-0.307	-0.056	1.246	-0.208	0.615	0.044
13	-1.242	0.748	0.146	0.715	0.484	0.485	-1.159	0.788	0.289	-1.521
14	-0.481	-0.269	-1.421	0.949	-1.111	-0.463	-0.406	0.585	-0.897	0.870
15	1.411	1.925	-0.771	0.417	-0.041	2.017*	-0.978	-0.342	0.429	-1.073

Critical values for the 0.05 probabillty level are ±1.96.

TABLE 4.31

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 50 SEQUENCES

GENERATED BY EQUATION (2.2.1)

Seq. No. / Lag p	1	2	3	4	5	6	7	8	9	10
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	1.454	1.359	-1.299	0.258	-0.074	-0.888	0.489	0.603	-0.373	-1.052
2	0.711	-0.927	-0.091	0.585	-0.485	-0.988	1.626	-0.239	-0.811	0.452
3	1.332	0.941	0.352	-0.525	0.448	-2.552*	-1.699	-1.544	-0.375	0.160
4	-0.344	-1.092	-1.210	-1.946	0.918	-0.544	-0.017	-0.471	1.204	-2.827*
5	0.484	1.033	1.047	0.254	-0.413	-0.472	1.304	-0.102	1.974*	-0.833
6	-1.529	1.385	-1.495	0.666	-1.270	-1.110	0.406	-1.223	-0.506	0.437
7	-1.446	-1.275	-0.205	1.644	0.476	-0.251	1.963*	0.925	0.320	-1.052
8	1.442	0.460	1.145	-0.882	-0.650	-2.374*	0.583	1.645	1.000	-0.850
9	0.897	-0.017	-0.338	-0.047	-0.602	1.802	0.002	-0.515	0.548	0.336
10	-0.518	-0.347	-0.621	-0.979	-0.741	-0.206	0.380	0.009	-0.726	0.122
11	0.741	0.418	0.989	2.127*	-1.033	0.280	0.173	-0.095	-0.578	0.721
12	0.444	0.081	-1.078	0.687	-2.894*	0.702	1.402	-0.442	-1.802	-0.027
13	-0.767	-0.106	0.427	0.065	0.849	0.768	0.263	0.090	0.164	-0.565
14	0.424	-0.576	1.466	-0.657	0.142	0.968	-0.778	-0.261	1.262	0.539
15	1.361	-0.684	-0.088	-0.156	0.427	-0.444	0.083	-0.718	-0.958	0.827

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.31

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 50 SEQUENCES

GENERATED BY EQUATION (2.2.1) - Continued

Seq. No. / Lag p	11	12	13	14	15	16	17	18	19	20
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-1.333	0.144	0.779	-0.251	-1.682	0.386	0.168	0.950	0.854	1.551
2	0.310	1.798	0.420	1.626	0.670	0.399	-0.945	0.030	-0.259	-0.129
3	-0.840	-1.359	-1.812	1.278	1.660	0.527	-0.296	0.038	-0.017	-0.001
4	-0.896	-1.217	0.361	-2.518*	-0.131	-1.415	0.559	0.254	-0.808	1.168
5	-0.726	0.254	-0.304	0.342	-0.631	-0.333	0.947	-0.706	-1.352	2.248*
6	-0.685	0.939	0.339	0.202	0.195	-1.285	-1.028	0.354	0.539	-0.167
7	1.087	-0.004	-0.720	-1.477	0.896	1.319	1.230	1.465	-0.456	1.135
8	-1.562	-0.317	0.388	-1.412	-0.241	1.890	-0.826	2.544*	1.040	1.523
9	-0.717	-0.507	-0.964	1.628	0.020	-1.585	-1.517	-0.478	0.032	-0.502
10	0.169	1.625	0.178	-0.111	1.434	-0.065	1.750	2.023*	-0.251	2.229*
11	0.065	-1.324	-0.236	-0.421	0.557	1.141	0.636	0.414	-1.098	-0.584
12	0.878	-0.545	-0.052	0.187	2.089*	0.555	0.521	-0.673	-0.585	-0.777
13	0.235	-0.223	0.172	-1.117	-0.128	1.439	1.226	0.697	-0.318	0.642
14	-1.107	-0.325	-1.618	0.718	-1.051	1.271	0.088	-0.622	0.051	-0.170
15	-0.815	2.579*	-0.769	0.793	-1.439	1.332	1.062	0.427	0.407	-0.732

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.31

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 50 SEQUENCES

GENERATED BY EQUATION (2.2.1) - Continued

Seq. No. / Lag p	21	22	23	24	25	26	27	28	29	30
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	0.098	0.011	-0.855	0.494	0.380	-0.247	1.274	1.401	0.996	-0.845
2	-1.162	-0.532	-0.516	-2.580*	-0.889	-1.006	-0.679	0.961	-0.744	-0.692
3	1.052	-0.608	0.746	-0.373	-0.680	-0.325	-0.241	-0.900	-0.252	-0.339
4	-0.620	-1.902	0.122	-0.728	1.217	-1.522	-0.620	0.937	-0.732	-1.535
5	0.237	1.481	-0.548	0.357	0.328	-0.368	0.711	0.571	-1.282	-0.197
6	-0.353	0.084	-0.877	-2.493*	1.005	0.529	1.200	0.140	2.148*	-1.604
7	-0.104	0.127	-0.860	-1.509	-0.467	-0.969	1.494	0.194	-1.305	0.451
8	1.177	-0.668	-0.181	-0.235	-1.291	0.320	-0.185	-1.343	-0.943	-2.097
9	-0.457	0.238	-0.507	0.293	0.450	-0.782	-1.651	0.029	2.311*	1.240
10	-2.982*	0.337	-0.403	-0.996	1.738	0.772	-0.056	0.376	0.154	-0.318
11	0.924	0.296	-0.370	0.567	-0.986	-0.417	1.396	-1.695	1.659	-0.288
12	0.824	0.768	0.351	-0.034	0.823	0.632	0.723	-0.323	0.927	-0.139
13	1.553	-1.242	0.150	1.097	1.099	0.709	0.423	1.362	0.235	0.659
14	-0.197	-0.066	0.303	-0.073	1.362	-1.053	-0.359	-0.155	0.946	0.627
15	1.489	1.629	-1.104	1.604	-0.132	1.070	1.542	-0.467	-0.977	0.030

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.31

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 50 SEQUENCES

GENERATED BY EQUATION (2.2.1) - Continued

Lag p \ Seq. No.	31	32	33	34	35	36	37	38	39	40
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-2.998*	-1.379	1.337	1.260	-0.006	0.340	-0.496	-0.856	-0.494	-0.827
2	0.176	-0.063	-0.648	0.629	0.259	-1.166	-0.028	-0.270	0.311	0.127
3	0.016	-1.281	0.437	-0.505	0.813	-0.721	-0.472	-1.383	-1.250	0.076
4	-1.047	-0.954	0.062	-1.213	-0.416	1.402	-0.358	-1.192	-1.963*	0.270
5	0.227	1.711	1.264	-1.348	-1.076	1.480	-0.720	-1.298	-0.181	0.334
6	2.484*	0.377	-0.016	0.472	-1.112	-0.814	-0.297	2.230*	0.315	1.497
7	-0.851	-1.316	3.233*	-0.302	1.832	-0.275	0.217	0.202	-0.412	-0.616
8	-1.166	-0.996	-0.813	-0.646	-0.329	0.690	-0.894	0.301	0.110	-0.346
9	-0.257	0.445	-0.714	-1.022	-0.884	-0.349	0.226	0.986	0.693	0.933
10	1.054	-1.262	-0.345	0.123	0.714	-1.473	0.471	-0.579	-0.470	-0.511
11	-1.717	0.339	-0.369	0.318	0.976	-0.399	-0.805	0.468	0.404	0.072
12	0.284	-0.637	0.697	0.884	0.674	-1.585	-0.175	0.243	-0.549	1.216
13	1.226	0.714	-1.306	-0.188	0.665	0.517	0.099	0.730	0.013	-0.525
14	-0.543	-1.552	0.487	1.207	-1.531	1.257	1.183	0.502	1.327	0.558
15	-0.595	-1.281	-1.490	2.449*	0.590	0.226	2.322*	-0.217	0.376	-0.114

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.31

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 50 SEQUENCES

GENERATED BY EQUATION (2.2.1) - Concluded

Seq. No. / Lag p	41	42	43	44	45	46	47	48	49	50
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	0.667	-0.741	0.573	-0.830	-1.610	0.165	0.825	-0.233	-1.002	-0.578
2	-0.038	0.930	0.095	-0.990	0.909	0.405	-2.330*	0.570	-0.124	-0.460
3	-0.019	-0.174	-0.395	0.348	0.429	0.850	1.653	0.278	-0.026	0.251
4	0.409	-0.090	1.093	0.859	0.476	-0.971	0.656	-0.534	0.586	0.190
5	-0.210	-0.823	0.121	-2.258*	0.370	-0.231	-0.686	1.309	-0.613	-0.367
6	1.251	1.305	0.113	0.098	-0.880	1.728	1.029	0.794	-0.126	1.290
7	-0.285	-0.084	2.059*	-0.743	-1.053	-0.216	-1.372	0.772	0.232	-0.658
8	-2.018*	0.074	0.048	-1.239	0.992	0.410	1.804	1.258	0.033	-0.790
9	-0.403	0.465	0.905	-0.792	0.364	-0.599	0.560	-0.857	-1.611	1.201
10	-0.151	1.282	0.643	-0.359	0.244	0.017	-0.940	1.376	-0.698	0.660
11	-1.370	-1.997*	-0.622	-1.781	2.191*	-1.629	1.408	1.972*	-0.693	-0.256
12	0.565	-0.572	1.724	0.189	3.376*	-0.307	1.168	-2.416*	1.282	0.960
13	1.808	-0.982	0.351	0.816	0.990	-0.031	-0.504	-0.799	0.263	0.361
14	0.320	0.699	-1.258	0.037	-0.104	-0.862	-0.112	-0.968	0.260	1.977*
15	0.259	2.112*	0.322	0.558	-0.234	0.253	1.190	-0.095	0.077	-1.001

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.32

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 50 SEQUENCES

GENERATED BY EQUATION (2.2.2)

Lag p \ Seq. No.	1	2	3	4	5	6	7	8	9	10
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-2.904*	-2.182*	-2.762*	-1.710	-2.483*	-2.233*	-2.687*	-1.336	-1.772	-1.877
2	-1.301	-0.390	-0.147	0.342	-0.926	-0.042	-0.040	-0.573	-0.172	0.900
3	0.216	0.614	1.065	-0.640	-0.153	-0.879	0.395	-1.223	0.177	2.158*
4	-0.511	-0.652	1.104	2.697*	-0.773	0.901	-1.800	-0.380	-0.495	-0.802
5	0.348	1.691	2.148*	1.000	2.551*	-0.338	-0.401	1.193	-1.569	-2.419*
6	-0.895	0.200	0.428	-0.284	-0.218	-0.932	0.950	-0.924	-1.183	-0.714
7	0.957	-2.477*	0.328	-0.145	-0.481	0.762	0.883	1.557	-1.264	-0.276
8	-0.951	0.736	0.180	0.948	-0.404	-0.083	0.227	-0.901	0.372	-1.619
9	-0.036	-0.671	0.808	-0.368	0.702	-0.708	-0.121	-0.126	-0.042	0.064
10	-1.287	1.011	-0.502	-0.361	0.474	-0.194	-2.964*	0.938	-0.929	-2.021*
11	0.665	-0.651	-0.038	0.639	1.130	-1.016	0.720	-0.200	0.397	0.577
12	-0.640	0.139	-0.818	1.133	-0.938	-0.634	0.585	-1.069	-1.049	-0.708
13	1.358	1.103	0.462	1.096	0.670	-0.573	0.175	-1.488	1.126	-1.383
14	-0.558	0.310	-1.098	-0.536	-0.387	-0.420	0.719	0.115	2.463*	-0.432
15	-0.150	-0.310	-1.688	0.319	0.671	0.795	0.496	2.061*	-0.042	1.150

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.32

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 50 SEQUENCES

GENERATED BY EQUATION (2.2.2) - Continued

Seq. No. / Lag p	11	12	13	14	15	16	17	18	19	20
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-2.246*	-4.079*	-1.968*	-4.506*	-4.527*	-2.323*	-4.169*	-3.156*	-3.109*	-2.812*
2	0.095	-0.535	-0.933	-0.702	-0.326	-0.410	0.013	-0.087	0.012	0.018
3	-0.066	-1.207	-1.013	-0.624	-0.236	1.152	0.535	1.808	1.039	0.130
4	-0.055	1.644	-0.424	-0.336	-0.953	1.583	0.648	-1.302	-0.959	0.302
5	0.962	-0.589	0.443	1.452	0.195	1.612	0.099	-0.013	0.004	-0.681
6	0.189	-0.706	-1.094	1.461	-0.348	-0.618	1.432	0.221	-0787	-0.962
7	-0.258	0.398	-0.902	-0.369	1.488	0.290	1.226	-0.931	-0.052	0.964
8	-1.960*	-0.792	-0.325	-0.300	-1.875	0.098	-0.272	1.583	-0.443	-0.125
9	-0.911	1.626	-1.504	-0.602	0.384	0.047	0.968	-1.173	1.370	-1.062
10	-0.858	-1.239	1.524	-0.641	-0.541	-0.167	-0.365	-2.347*	-0.693	-0.581
11	-1.121	-1.095	-0.265	-1.098	1.723	-1.011	-0.994	1.416	-0.050	-0.783
12	-0.056	1.480	-0.458	-0.267	-0.053	0.026	0.052	-0.493	-0.114	-0.654
13	0.864	-0.540	-0.135	1.523	-0.822	-0.175	-0.049	-0.092	0.161	-1.552
14	-1.654	-1.108	-0.217	-2.036*	0.117	-0.476	-0.620	1.657	3.305*	-0.055
15	-1.421	-0.354	1.051	0.739	0.168	-0.984	1.634	0.155	0.118	0.648

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.32

RESULTS OF THE TEST FOR CORRELATION FOR 15 LAGS OF 50 SEQUENCES

GENERATED BY EQUATION (2.2.2) - Continued

Seq. No. / Lag p	21	22	23	24	25	26	27	28	29	30
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-1.431	-3.172*	-0.562	-3.021*	-3.640*	-3.344*	-2.850*	-2.376*	-3.361*	-3.577*
2	-1.635	-1.241	-1.639	0.006	-0.845	-1.758	0.064	0.177	-0.713	-0.415
3	0.480	0.170	0.376	-0.621	1.722	0.155	-0.869	-0.241	0.475	0.483
4	-0.519	0.432	-0.231	-0.792	-0.946	-0.305	-0.566	0.141	0.328	-1.144
5	0.836	-0.038	-1.099	1.363	1.435	-1.095	0.263	0.902	1.916	0.282
6	0.013	1.243	-2.457*	0.755	0.712	-1.239	-0.462	0.598	0.486	0.103
7	1.694	0.175	1.703	1.324	-2.663*	-0.865	-0.330	-0.806	0.603	-2.570*
8	-0.276	0.203	1.511	-1.787	0.918	-0.317	-0.229	-1.005	-0.318	1.298
9	0.350	1.293	-0.301	-0.244	0.091	-1.652	-1.993*	0.189	0.900	0.540
10	-0.203	0.059	-0.429	-0.026	0.743	1.283	1.183	-1.689	0.487	0.676
11	0.296	-1.076	0.315	1.566	-1.129	2.104*	-0.954	-0.130	0.622	0.473
12	1.070	-0.102	-0.283	0.710	1.268	-0.138	1.091	-0.416	-0.787	0.772
13	-0.298	1.319	-0.378	-0.480	-0.931	1.392	0.301	0.207	-0.788	-0.347
14	-0.307	-1.821	-1.408	2.185*	0.486	-2.067*	0.720	-2.198*	0.406	-1.397
15	0.886	1.326	-0.801	0.051	1.374	0.854	0.190	-0.436	-0.023	-0.047

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.32

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 50 SEQUENCES

GENERATED BY EQUATION (2.2.2) - Continued

Seq. No. / Lag p	31	32	33	34	35	36	37	38	39	40
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	1.781	-3.032*	1.415	-1.688	-4.399*	-3.369*	-2.980*	-2.883*	-2.042*	-3.552*
2	-0.723	-0.346	1.127	0.848	0.200	-1.242	0.566	-1.689	0.934	-0.305
3	0.088	-1.468	1.522	0.422	-0.383	0.063	-1.017	0.706	0.570	0.506
4	0.563	0.205	-0.062	-0.308	-1.069	0.105	-0.010	-0.255	-0.233	-1.052
5	-0.405	-0.748	-0.780	-1.544	-0.550	0.964	-0.378	0.028	-1.162	0.695
6	-1.253	-1.500	-0.692	0.932	1.470	0.120	1.070	-0.752	-0.478	1.160
7	-1.255	1.377	-1.444	-1.005	-0.265	0.490	0.827	0.552	-0.379	0.415
8	-0.685	-0.704	1.537	0.036	0.275	2.818*	-0.452	2.721*	0.087	-1.153
9	0.244	0.068	0.837	1.123	0.818	-1.208	0.221	-1.580	1.304	0.255
10	-0.857	-0.937	-0.276	-1.099	1.231	-1.038	0.659	-1.354	0.110	0.265
11	1.521	2.154*	-0.575	-0.548	-0.708	-0.029	0.679	-1.126	0.135	2.213*
12	-0.273	1.845	0.860	1.440	-1.083	1.631	1.450	0.952	-0.161	0.208
13	-1.046	0.575	-0.351	0.517	1.216	0.470	0.179	-1.453	-0.112	-0.719
14	0.615	-0.259	-0.503	0.067	0.497	0.288	-0.193	1.249	-0.015	0.937
15	-0.297	0.022	-0.536	-0.531	-0.157	-0.034	0.133	0.845	-0.952	1.619

Critical values for the 0.05 probability level are ±1.96.

TABLE 4.32

RESULTS OF THE TEST FOR ZERO CORRELATION FOR 15 LAGS OF 50 SEQUENCES

GENERATED BY EQUATION (2.2.2) - Concluded

Seq. No. / Lag p	41	42	43	44	45	46	47	48	49	50
	$Z = \frac{R - E(R)}{\sigma_R}$									
1	-2.012*	-4.513*	-2.718*	-3.110*	-2.309*	-2.628*	-3.553*	-3.599*	-3.894*	-2.557*
2	0.663	-2.182*	-0.505	0.613	0.668	-1.496	-1.072	-1.226	-2.745*	-0.270
3	-0.470	0.202	0.161	1.275	0.134	-1.853	-0.220	2.949*	3.117*	0.425
4	-0.551	0.511	-0.417	0.548	1.303	0.333	0.281	-0.006	-1.207	-0.553
5	-0.014	-0.885	-0.018	-1.918	0.492	0.477	-0.742	-1.412	0.260	0.363
6	0.163	0.649	0.156	0.771	-0.622	1.597	1.130	0.061	-0.755	-0.040
7	1.741	-0.119	-3.043*	-1.230	0.058	-1.167	-0.678	0.722	0.736	1.420
8	-0.099	-0.341	0.260	0.922	-0.291	0.655	0.001	-0.743	-0.341	-0.504
9	-0.601	1.365	-0.638	0.352	0.368	1.626	1.069	1.483	0.578	-0.357
10	0.048	0.040	-0.338	-0.736	-1.103	-0.863	-0.051	0.334	-0.040	-1.411
11	-0.639	-1.641	-0.440	1.079	-1.215	1.394	-0.962	0.195	-0.765	0.050
12	0.496	-0.499	-0.050	0.259	-0.121	-2.528*	-0.442	-1.122	-1.345	1.232
13	-0.678	1.363	0.191	1.534	-0.762	-0.504	1.816	0.586	-0.300	0.148
14	0.873	-0.349	0.059	1.108	-0.759	0.618	-1.285	1.758	1.131	0.895
15	0.380	0.773	1.779	0.839	0.086	0.200	0.332	-0.691	-0.627	0.588

Critical values for the 0.05 probability level are ±1.96.

V. TESTS FOR NORMALITY AND UNIFORMITY

Since for each sequence of numbers generated by the methods in Chapter II a distribution function is specified, it would be reasonable to expect each sequence of numbers to possess all the essential properties of the specified distribution function. One's interest in this chapter is somewhat different than in the previous one. Here one is concerned mainly on whether or not each sequence of numbers truly comes from the specified distribution. A sequence of numbers may be random; however, if it does not follow a prescribed distribution, it could be useless.

The tests that follow should indicate with some degree of confidence whether each sequence follows the uniform distribution on the internal (0,1) or the normal distribution with zero mean and unit variance. For this purpose, 4 tests have been performed on each of the 200 sequences generated by the methods of Chapter II for this study. Two of the tests are Karl Pearson's χ^2 and the Kolmogorov-Smirnov D-test statistic. Both of these are known as goodness-of-fit tests. The remaining two are tests on the sequence means and variances, respectively.

5.1 Chi-Square Goodness-of-Fit Test

Let $x_1, x_2, \ldots, x_N$ be N independent observations of some random variable with an unknown distribution function $f(X)$. Consider testing the null hypothesis

$$H_0: \ f(X) = f_0(X) \tag{5.1.1}$$

where $f_0(X)$ is some specified distribution function. In general, the problem of testing such a hypothesis is known as "goodness-of-fit" problem. If $f_0(X)$ is completely specified with regard to its parameters, (5.1.1) is referred to as a simple hypothesis. This case will be assumed here.

Let $x_1, x_2, \ldots, x_N$ be the N observations of a sequence of numbers generated by the methods in Chapter II. Consider the division of these N observations into k mutually exclusive classes each class containing N_i observations. Let the probability of an observation falling in the ith class be

$$p_i = \frac{N_i}{N} \qquad i = 1, 2, \ldots, k \tag{5.1.2}$$

where $\sum_{i=1}^{k} p_i = 1$. Since it is assumed that $f_0(X)$ is completely specified, the corresponding class probabilities according to the k mutually exclusive classes may be determined and denoted by p_{io} where, necessarily, $\sum_{i=1}^{k} p_{io} = 1$. Consider the null hypothesis (5.1.1) to be started as follows:

$$\begin{aligned} H_0&: \ p_i = p_{io} \\ & \qquad\qquad\qquad i = 1, 2, \ldots, k \\ H_A&: \ p_i \neq p_{io} \end{aligned} \tag{5.1.3}$$

Then, when H_0 is true and from the k mutually exclusive classes, the test statistic

$$\tau = \sum_{i=1}^{k} \frac{(N_i - Np_{io})^2}{Np_{io}} \qquad (5.1.4)$$

is distributed approximately as a chi-square with $k - 1$ degrees of freedom. The null hypothesis is rejected whenever the computed τ exceeds the upper tail of the chi-square distribution with $k - 1$ degrees of freedom and some given level of significance. This is the well-known Karl Pearson's chi-square test of goodness-of-fit (16).

It has been shown by many (17) that this test statistic is asymptotically equivalent to the maximum likelihood ratio test. Much discussion has centered around the optimum choice of k for a given number of N observations. Generally, no power of this test statistic has yet been determined. Mann and Wald (18) have shown that if

$$k = 4\sqrt[5]{\frac{2(N - 1)^2}{c^2}} \qquad (5.1.5)$$

where N is the number of observations in a given sequence and c is determined so that $\frac{1}{\sqrt{2\pi}} \int_c^{\infty} e^{-\frac{1}{2}t^2}\, dt$ is equal to the probability of the critical region under the null hypothesis (5.1.3), then the power of the test statistic is approximately 0.50. For a large N, (5.1.5) would produce a fairly large k which could be impractical. However, Williams (19) and Kendall and Stuart (17) report that k can be halved from (5.1.5) without serious loss of power at the 0.50 point. This suggestion is considered in this study. It is worth noting that (5.1.5) would be an effective measure of k whenever N is substantially large.

5.1.1 Pearson's Test for Normality

To apply Pearson's chi-square test to each sequence generated by equations (2.2.1) and (2.2.2), one must classify the $N = 10,000$ observations of each sequence into some k classes. For $N = 10,000$ and an α equal to 0.05 for the probability of Type I error, one determines a k by (5.1.5) to be approximately equal to 140. Recall that

$$X_1 = (-2 \ln U_1)^{1/2} \sin 2\pi U_2$$

$$X_2 = (-2 \ln U_1)^{1/2} \cos 2\pi U_2$$

were a pair of independent random variables having the same normal distribution with mean zero and unit variance. Consider the following class arrangement of a given sequence generated by one of the above equations:

$$(-\infty, -3.0), (-3.0, -2.9), (-2.9, -2.8) \; . \; . \; . \; (2.9, 3.0), (3.0, \infty)$$

This arrangement would produce $k = 62$ classes of 0.10 width. This is fairly close to one-half the value of k obtained by (5.1.5). For this classification, the observed probabilities, p_i, may be determined by dividing the number observed in each class, N_i, by N for $i = 1, 2, \ldots, 62$. The corresponding theoretical probabilities, p_{io}, may be obtained from any table of the standard normal distribution. Thus, to test the null hypothesis

$$\begin{aligned} H_0 &: \; p_i = p_{io} \\ & \qquad\qquad\qquad i = 1, 2, \ldots, 62 \\ H_A &: \; p_i \neq p_{io} \end{aligned} \tag{5.1.2}$$

compute the statistic

$$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}} \tag{5.1.3}$$

and compare it with the upper tail of the chi-square distribution with $k - 1 = 61$ degrees of freedom and a probability of 0.05 for Type I error. The null hypothesis is rejected if the computed τ exceeds the critical region. Rejection implies that the distribution of the given sequence of N observations is something other than the normal with zero mean and unit variance.

The results of this test on the 100 sequences generated by equations (2.2.1) and (2.2.2) appear systematically in tables (5.10) and (5.11). Out of the possible 100 sequences, 4 were rejected. Hence, the results indicate that both equations (2.2.1) and (2.2.2) generate sequences that appear to be normally distributed with zero mean and unit variance.

5.1.2 Pearson's Test for Uniformity

To test that a sequence of numbers generated by equation (2.1.4) is uniformly distributed on the unit interval, one uses a similar approach as in the case with normality. Consider the classification of the $N = 10{,}000$ observations for each sequence to be the following:

(0.00, 0.01), (0.01, 0.02), (0.02, 0.03) . . . (0.98, 0.99), (0.99, 1.00)

This results into $k = 100$ classes of 0.01 width. It is apparent that this is a natural division of the unit interval in such a way as to produce 100 classes with equal probabilities. Let N_i be the number of observations falling in the ith class. Clearly, the corresponding theoretical class probabilities, p_{io}, equal to 0.01 for all $i = 1, 2, \ldots, 100$. Hence to test the null hypothesis of uniformity, compute the statistic

$$\tau = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N} \tag{5.1.4}$$

and compare the computed τ with the upper tail of the chi-square distribution with 99 degrees of freedom and $\alpha = 0.05$. Rejection implies that the distribution of the given sequence is not the uniform on the interval (0,1).

Table (5.12) contains the results of this test statistic for the 100 sequences generated by equation (2.1.4). There was a total of three sequences rejected. Therefore, the apparent conclusion is that equation (2.1.4) seems to generate sequences uniformly distributed on the unit interval.

5.2 Kolmogorov-Smirnov Criterion

Consider the cumulative distribution function $F_O(X)$ of some specified density function $f(X)$. Clearly, for any specified value of the random variable X, the value of $F_O(X)$ is the proportion of individuals in the population having values less than or equal to the specified X. Consider, further, a sequence of N observations thought of having the distribution - and hence the cumulative $F_O(X)$ - $f(X)$. It is quite reasonable to expect the cumulative step-function of the sequence of N observations to be fairly close to the specified cumulative distribution function $F_O(X)$. If this is not the case, one may reasonably assume that $f(X)$ is not the distribution function of the sequence.

Let $F_O(X)$ be the cumulative distribution function of the specified density function $f(X)$, and $F_N(X)$ the observed cumulative step-function of the sequence of N observations. That is, $F_N(X) = k/N$ where k is the number of observations less than or equal to a specified X. Then the distribution of

$$D = \max\left|F_N(X) - F_O(X)\right| \tag{5.2.1}$$

is known and is independent of $F_O(X)$ if $F_O(X)$ is continuous. The limiting distribution of this test statistic D has been derived by Kolmogorov (13) and by Smirnov (14); thus, it is known as the Kolmogorov-Smirnov goodness-of-fit test. Hence, to test the null hypothesis that a given sequence of N observations has the specified cumulative distribution function $F_0(X)$, record in a table the observed and specified cumulative distribution functions and calculate the maximum deviation between them. Reject the null hypothesis whenever the maximum deviation exceeds the critical value of the D-statistic with some α, the probability of Type I error.

This test becomes exact whenever the observations are not grouped into various classes. However, for most applications this is prohibitive. Although grouping observations into class intervals tends to lower the value of D, for large samples, the discrepancy is usually very slight and will cause little change in the appropriate significance levels. In addition, some caution must be exercised not to choose a very small number of class intervals.

It was noted in the preceding section that in general no power of Pearson's X^2 goodness-of-fit test exists. However, a lower bound of the power of the D-statistic has been derived by Massey (11). The results of his work are presented below. Let $F_1(X)$ be an alternative form of the specified cumulative distribution function $F_0(X)$. Let Δ be the maximum absolute difference between $F_1(X)$ and $F_0(X)$. Then for N large, it has been shown (15) that the power of the D-statistic is never less than

$$1 - \frac{1}{\sqrt{2\pi}} \int_{2\left[\Delta\sqrt{N} - D_\alpha(N)\right]}^{2\left[\Delta\sqrt{N} + D_\alpha(N)\right]} e^{-t^2/2}\, dt \tag{5.2.2}$$

where N is the number of observations and $D_\alpha(N)$ is the critical value of the D-statistic with α being the probability of Type I error. Based on the work of Massey (11), (15), figure 5.2 shows this lower bound for the 5 and 1 percent levels of significance. Hence, the figure indicates that the D-statistic of $F_0(X)$ has at least 0.50 power against the alternative $F_1(X)$. As an example, suppose $N = 10{,}000$, and the maximum absolute difference between $F_1(X)$ and $F_0(X)$ is 0.02; then, $\Delta\sqrt{N} = 2.00$. If one tests the specified distribution $F_0(X)$ at the 5 percent level of significance, a lower bound of the power of this test is found to be 0.89. Therefore, the power of this test is at least 0.89 against the alternative $F_1(X)$. In other words, if $F_1(X)$ is correct, one has at least 89 percent chance of detecting that $F_0(X)$ is the incorrect cumulative distribution function of the given sequence of $N = 10{,}000$ observations.

The author has considered both existing tests of goodness-of-fit; namely, Pearson's χ^2 and the Kolmogorov-Smirnov D-statistic. In general, the power of Pearson's χ^2 is not known, where a lower bound to the power of the D-statistic may be read from figure 5.2 for any alternative. Judging on this comparison alone, it is reasonable to believe that the D-statistic provides the better test for goodness-of-fit as long as the specified distribution function is continuous.

5.2.1 Kolmogorov-Smirnov D-Statistic for Normality

To test any sequence of numbers generated by equations (2.2.1) or (2.2.2) for normality using the D-statistic, classify the observations

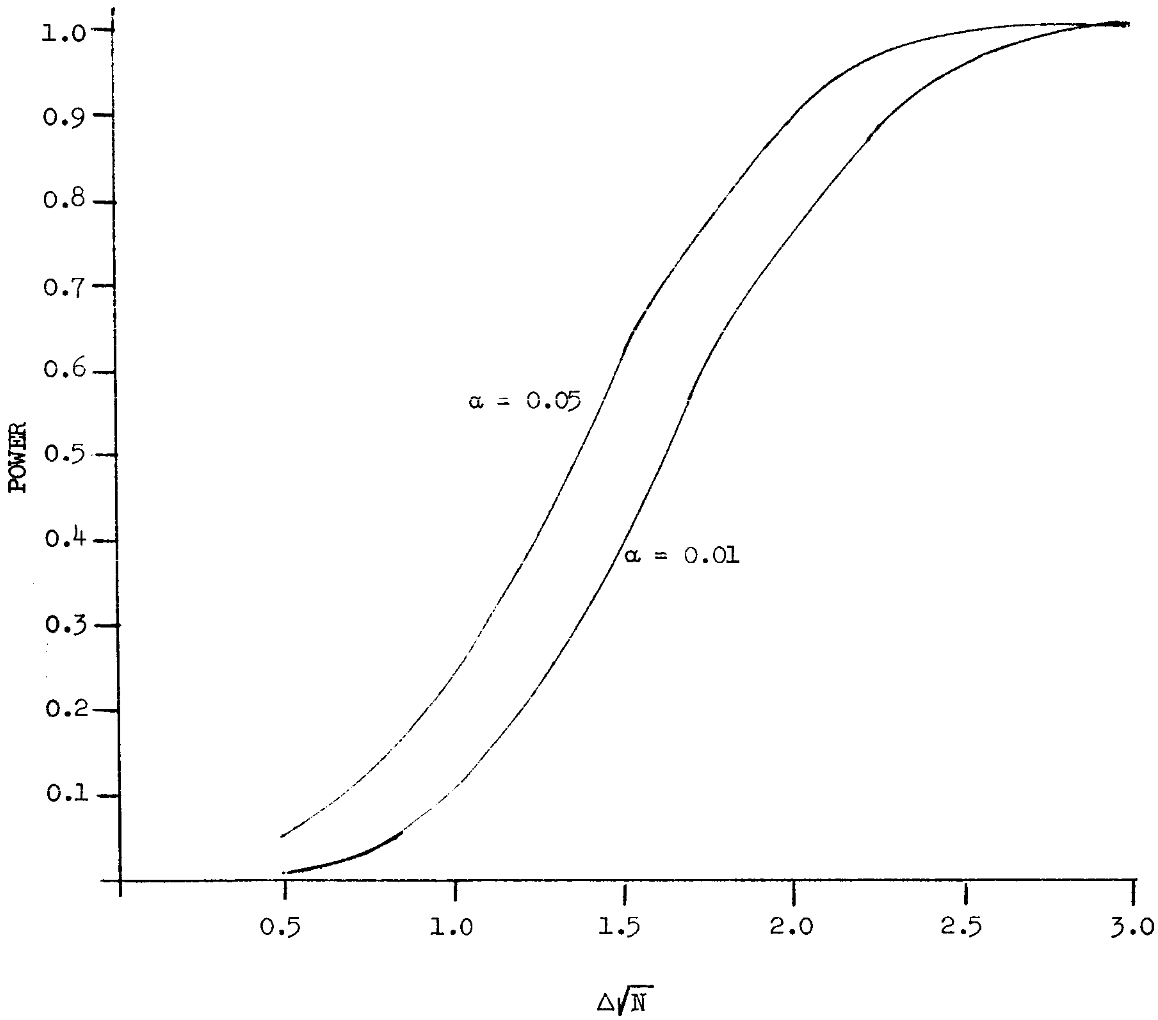

FIGURE 5.2

Lower bounds for the power of the D test for $\alpha = 0.01$ and $\alpha = 0.05$
(Reproduced from (11))

into the 62 class intervals of 0.10 width as indicated in 5.1.1. Obtain the cumulative distribution function $F_0(X)$ for the specified normal density and the above class intervals from any appropriate table of the standard cumulative normal distribution function. Determine the observed cumulative step-function $F_N(X)$ by dividing the number of observations in each class interval by $N = 10,000$. Compute the maximum absolute difference

$$D = \max \left| F_N(X) - F(X) \right| \tag{5.2.2}$$

and compare it with the critical value $D_\alpha(N)$, where $\alpha = 0.05$ and $D_{0.05}(10,000)$ is given by (11) $\frac{1.36}{\sqrt{N}} = \frac{1.36}{\sqrt{10,000}} = 0.0136$. Reject the null hypothesis of normality if the computed D exceeds $D_{0.05}(10,000)$. The D-statistic has been performed on each of the 50 sequences generated by equation (2.2.1) and the 50 sequences generated by equation (2.2.2). The results appear in tables (5.20) and (5.21), respectively. These indicate that both equations (2.2.1) and (2.2.2) seem to be generating sequences having the specified normal distribution. Although no rejections were noted for any of the 100 sequences, this should not be an alarming result because from the discussion of the preceding section a lower bound of the power of the D-statistic can be easily determined. For example, even if the maximum absolute difference between the specified $F_0(X)$ and some alternative $F_1(X)$ is as small as 0.0175, one has at least 78 percent chance of detecting the incorrectness of $F_0(X)$ at the 5 percent level.

5.2.2 Kolmogorov-Smirnov D-Statistic for Uniformity

A very similar procedure may be followed to test the null hypothesis that a sequence of numbers generated by (2.1.4) is uniformly distributed on the interval (0,1) by considering the 100 class intervals of 0.01 width as indicated in 5.1.2. By determining the specified cumulative distribution function $F_0(X)$, and the observed cumulative step-function $F_N(X)$ based on the 100 class intervals, the statistic

$$D = \max \left| F_N(X) - F_0(X) \right| \qquad (5.2.3)$$

may be computed and compared with $D_{0.05}(10{,}000) = 0.0135$ as before. Rejection implies that the distribution of the given sequence is not the uniform on the unit interval. The results of this test on the 100 sequences generated by equation (2.1.4) may be found in table (5.22). The results indicate the following:

A total of four rejections were noted; this, of course, is within the framework of the statistical test. Therefore, it appears that equation (2.1.4) does generate sequences that are uniformly distributed on the interval (0,1).

5.3 Tests on Sequence Means and Variances

Since each sequence of numbers is assumed to be either uniformly or normally distributed - depending on whichever is the case - with some specified mean and variance, it is reasonable to expect each mean and variance of a given sequence to estimate or be approximately the same as the mean and variance of the specified distribution function.

To detect the degree of agreement, the following common statistical tests have been performed on each sequence.

5.3.1 Tests on Sequence Means

Let $\overline{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$ be the mean of a given sequence $x_1, x_2, \ldots,$ x_N. Then, for any given distribution function with mean μ and variance σ^2, $\overline{x}$ is normally distributed with mean μ and variance σ^2/N, and the statistic

$$Z = \frac{\overline{x} - \mu}{\sigma/\sqrt{N}} \tag{5.3.1}$$

is the value of a random variable whose distribution function approaches that of the standard normal as $N \rightarrow \infty$. This, of course, is the central limit theorem on which one bases a test to determine whether or not the mean of some sequence of numbers is truly the mean of the specified distribution function of that sequence.

Hence, to test the null hypothesis H_0: $\mu = \mu_0$ against H_A: $\mu \neq \mu_0$, where μ_0 is equal to the mean μ of the specified distribution, compute the statistic

$$Z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{N}} \tag{5.3.2}$$

and compare it to both the left and the right tails of the standard normal with a chosen probability of 0.05 for Type I error. If Z exceeds these limits, reject the null hypothesis and conclude that the mean of the sequence being tested is something other than the mean of the specified distribution.

Results of this test for the 200 sequences may be found in tables (5.30), (5.31), and (5.32). These results indicate the following:

(a) For the 100 sequences generated by equations (2.2.1) and (2.2.2), no rejections were noted. As a matter of fact, the agreement between sequence means and the theoretical mean of zero is excellent.

(b) For the 100 sequences generated by equation (2.1.4), a total of seven rejections were noted. However, on the average the agreement with the theoretical mean of 0.50 was very good.

5.3.2 Tests on Sequence Variances

Throughout this study, each sequence of numbers is assumed to have one of two distributions, the uniform or the normal. A test on the variance of a given sequence assumed to be normally distributed is quite straightforward. Let the unbiased estimator of σ^2

$$s^2 = \frac{\sum_{i=1}^{N} x_i^2 - \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{N}}{N - 1} \tag{5.3.3}$$

be the variance of a given sequence of numbers assumed to be normally distributed. To test the null hypothesis H_0: $\sigma^2 = \sigma_0^2$ against

H_A: $\sigma^2 \neq \sigma_0^2$ where σ_0^2 is equal to the variance σ^2 of the specified normal distribution, compute the statistic

$$\chi^2 = \frac{(N - 1)s^2}{\sigma_0^2} \tag{5.3.4}$$

where χ^2 is a value of a random variable having the chi-square distribution with $N - 1$ degrees of freedom. If $\chi^2 < \chi^2_{1-\alpha/2}$ or $\chi^2 > \chi^2_{\alpha/2}$ with $N - 1$ degrees of freedom and $\alpha = 0.05$, reject the null hypothesis and conclude that s^2 is not in agreement with σ^2, the variance of the specified distribution function.

No such exact test exists if the distribution of a sequence is not assumed to be the normal. Nevertheless, an approximate test on the sequence standard deviation s is frequently used and offers a fairly good approximation if the sample size is large. It is common knowledge that the sample variance s^2 is an unbiased estimator of σ^2; but the sample standard deviation s is not an unbiased estimator of σ. However, for large samples the bias is small, and it is common practice to estimate σ with s.

As a result, let $x_1, x_2, \ldots, x_N$ be a generated sequence of numbers with some specified distribution function. Define s^2 as in (5.3.3). Then, for a large sequence size and under fairly general conditions, the distribution of $s = \sqrt{s^2}$ can be approximated (10) closely with a normal having the mean σ and a standard deviation $\sigma/\sqrt{2N}$, where σ^2 is the variance of the specified distribution function.

Hence, to test the null hypothesis H_0: $\sigma = \sigma_0$ against H_A: $\sigma \neq \sigma_0$ where $\sigma_0 = \sqrt{\sigma^2}$, compute the statistic

$$Z = \frac{s - \sigma_0}{\sigma/\sqrt{2N}} \tag{5.3.5}$$

which is a value of a random variable having approximately the standard normal distribution. The critical values of Z may be determined as before.

This test has been performed on the 100 sequences generated by equation (2.1.4). The results for the entire 200 sequences may be found in tables (5.33), (5.34), and (5.35). The apparent conclusions are the following:

(a) For the 100 sequences generated by equations (2.2.1) and (2.2.2), there were six and five rejections, respectively. Although this number is slightly higher than expected, the agreement to the theoretical variance of unity is generally very good.

(b) For the 100 sequences generated by equation (2.1.4), no rejections were noted. This could be due to the fact that the test used was only an approximation. Nevertheless, the agreement to the theoretical variance of 0.083333 is excellent.

5.4 Numerical Results

Results of the various statistical tests described in this chapter appear systematically in the tables that follow. As before, all calculations were made with the use of the IBM 7094 computer; some of the results have been rounded for presentation. The critical values for the appropriate random variables are indicated in each table. As before, rejection is indicated by an asterisk.

TABLE 5.10

RESULTS OF PEARSON'S χ^2 GOODNESS-OF-FIT TEST FOR NORMALITY OF 50 SEQUENCES GENERATED BY EQUATION (2.2.1)

Sequence No.

	1	2	3	4	5	6	7	8	9	10
$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}}$	49.66	36.37	45.69	67.35	44.29	38.77	62.72	83.37*	80.60*	67.41

Sequence No.

	11	12	13	14	15	16	17	18	19	20
$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}}$	42.76	65.31	69.02	79.81	45.14	51.91	51.81	58.16	69.11	62.15

Sequence No.

	21	22	23	24	25	26	27	28	29	30
$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}}$	59.49	42.30	58.08	63.17	57.07	59.68	44.80	61.20	68.03	70.21

Sequence No.

	31	32	33	34	35	36	37	38	39	40
$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}}$	69.85	48.81	46.59	59.57	50.19	62.75	51.93	68.86	81.59*	55.25

Sequence No.

	41	42	43	44	45	46	47	48	49	50
$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}}$	69.86	44.98	57.70	69.02	69.69	43.43	70.23	72.24	48.85	48.21

The critical value of χ^2 with 61 degrees of freedom and the 0.05 probability level is 80.2.

TABLE 5.11

RESULTS OF PEARSON'S χ^2 GOODNESS-OF-FIT TEST FOR NORMALITY OF 50 SEQUENCES GENERATED BY EQUATION (2.2.2)

Sequence No.

	1	2	3	4	5	6	7	8	9	10
$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}}$	56.564	44.263	59.186	52.846	57.967	69.338	32.343	51.694	59.603	72.993

Sequence No.

	11	12	13	14	15	16	17	18	19	20
$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}}$	46.837	52.675	68.084	65.905	64.316	69.098	53.223	49.470	53.087	66.270

Sequence No.

	21	22	23	24	25	26	27	28	29	30
$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}}$	53.576	78.464	73.384	87.591*	51.955	64.107	59.948	73.764	46.105	50.364

Sequence No.

	31	32	33	34	35	36	37	38	39	40
$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}}$	46.052	65.304	59.543	59.506	52.193	55.302	62.191	52.041	60.632	63.082

Sequence No.

	41	42	43	44	45	46	47	48	49	50
$\tau = \sum_{i=1}^{62} \frac{(N_i - Np_{io})^2}{Np_{io}}$	51.042	69.498	44.166	52.677	50.418	60.914	58.833	69.999	64.023	51.671

The critical value of χ^2 with 61 degrees of freedom and the 0.05 probability level is 80.2.

TABLE 5.12

RESULTS OF PEARSON'S χ^2 GOODNESS-OF-FIT TEST FOR UNIFORMITY OF 100 SEQUENCES GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4)

Sequence No.

	1	2	3	4	5	6	7	8	9	10
$\tau = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N}$	96.22	90.26	93.80	68.00	81.06	130.48*	105.98	94.90	75.72	101.04

Sequence No.

	11	12	13	14	15	16	17	18	19	20
$\tau = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N}$	88.00	76.84	121.60	98.20	100.18	80.90	86.84	88.54	107.92	91.08

Sequence No.

	21	22	23	24	25	26	27	28	29	30
$\tau = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N}$	88.80	71.12	76.26	106.20	93.80	114.84	99.64	96.88	77.14	82.70

Sequence No.

	31	32	33	34	35	36	37	38	39	40
$\tau = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N}$	83.50	92.52	109.68	104.44	81.60	81.46	92.98	89.70	142.36*	92.18

Sequence No.

	41	42	43	44	45	46	47	48	49	50
$\tau = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N}$	100.74	110.00	72.98	95.24	75.10	72.40	80.58	107.70	127.74*	97.78

The critical value of χ^2 with 99 degrees of freedom and the 0.05 probability level is 123.2.

TABLE 5.12

RESULTS OF PEARSON'S χ^2 GOODNESS-OF-FIT TEST FOR UNIFORMITY OF 100 SEQUENCES GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Concluded

Sequence No.

	51	52	53	54	55	56	57	58	59	60
$T = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N}$	80.20	87.32	94.22	94.96	81.02	105.14	83.90	85.62	91.00	78.50

Sequence No.

	61	62	63	64	65	66	67	68	69	70
$T = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N}$	94.90	96.72	84.98	108.62	88.94	89.94	98.10	85.76	78.28	107.34

Sequence No.

	71	72	73	74	75	76	77	78	79	80
$T = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N}$	72.80	95.84	102.04	94.68	85.70	93.34	75.10	78.24	80.80	113.42

Sequence No.

	81	82	83	84	85	86	87	88	89	90
$T = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N}$	85.66	89.50	98.32	81.82	81.42	104.52	89.58	96.82	111.98	92.88

Sequence No.

	91	92	93	94	95	96	97	98	99	100
$T = \sum_{i=1}^{100} \frac{(N_i - 0.01N)^2}{0.01N}$	80.50	96.14	90.38	88.50	86.98	101.60	90.62	88.68	103.82	72.68

The critical value of χ^2 with 99 degrees of freedom and the 0.05 probability level is 123.2.

TABLE 5.20

RESULTS OF THE KOLMOGOROV-SMIRNOV GOODNESS-OF-FIT TEST FOR NORMALITY OF 50 SEQUENCES GENERATED BY EQUATION (2.2.1)

Sequence No.

	1	2	3	4	5	6	7	8	9	10
$D_N = \max\lvert F_N(X) - F_O(X)\rvert$	0.0053	0.0035	0.0056	0.0068	0.0031	0.0056	0.0069	0.0082	0.0083	0.0054

Sequence No.

	11	12	13	14	15	16	17	18	19	20
$D_N = \max\lvert F_N(X) - F_O(X)\rvert$	0.0053	0.0091	0.0047	0.0076	0.0070	0.0036	0.0108	0.0037	0.0102	0.0047

Sequence No.

	21	22	23	24	25	26	27	28	29	30
$D_N = \max\lvert F_N(X) - F_O(X)\rvert$	0.0103	0.0049	0.0050	0.0057	0.0082	0.0059	0.0056	0.0065	0.0057	0.0056

Sequence No.

	31	32	33	34	35	36	37	38	39	40
$D_N = \max\lvert F_N(X) - F_O(X)\rvert$	0.0054	0.0055	0.0060	0.0090	0.0080	0.0049	0.0066	0.0067	0.0076	0.0070

Sequence No.

	41	42	43	44	45	46	47	48	49	50
$D_N = \max\lvert F_N(X) - F_O(X)\rvert$	0.0097	0.0056	0.0063	0.0077	0.0101	0.0061	0.0061	0.0063	0.0086	0.0057

The critical value of D_k for $N = 10{,}000$ and the 0.05 probability level is 0.0136.

TABLE 5.21

RESULTS OF THE KOLMOGOROV-SMIRNOV GOODNESS-OF-FIT TEST FOR NORMALITY OF 50 SEQUENCES GENERATED BY EQUATION (2.2.2)

Sequence No.

	1	2	3	4	5	6	7	8	9	10
$D_N = \max\lvert F_N(X) - F_O(X)\rvert$	0.0053	0.0059	0.0039	0.0062	0.0104	0.0092	0.0053	0.0093	0.0064	0.0098

Sequence No.

	11	12	13	14	15	16	17	18	19	20
$D_N = \max\lvert F_N(X) - F_O(X)\rvert$	0.0039	0.0036	0.0039	0.0082	0.0060	0.0067	0.0084	0.0052	0.0108	0.0056

Sequence No.

	21	22	23	24	25	26	27	28	29	30
$D_N = \max\lvert F_N(X) - F_O(X)\rvert$	0.0079	0.0053	0.0099	0.0066	0.0052	0.0065	0.0049	0.0066	0.0051	0.0071

Sequence No.

	31	32	33	34	35	36	37	38	39	40
$D_N = \max\lvert F_N(X) - F_O(X)\rvert$	0.0077	0.0078	0.0072	0.0109	0.0068	0.0044	0.0059	0.0056	0.0067	0.0061

Sequence No.

	41	42	43	44	45	46	47	48	49	50
$D_N = \max\lvert F_N(X) - F_O(X)\rvert$	0.0050	0.0043	0.0081	0.0059	0.0079	0.0042	0.0049	0.0092	0.0088	0.0041

The critical value of D_K for $N = 10{,}000$ and the 0.05 probability level is 0.0136.

TABLE 5.22

RESULTS OF THE KOLMOGOROV-SMIRNOV GOODNESS-OF-FIT TEST FOR UNIFORMITY OF 100 SEQUENCES GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4)

Sequence No.

	1	2	3	4	5	6	7	8	9	10
$D_N = \max\lvert F_N(X) - F_0(X)\rvert$	0.0077	0.0040	0.0069	0.0055	0.0052	0.0195*	0.0072	0.0050	0.0087	0.0124

Sequence No.

	11	12	13	14	15	16	17	18	19	20
$D_N = \max\lvert F_N(X) - F_0(X)\rvert$	0.0070	0.0051	0.0078	0.0054	0.0094	0.0052	0.0077	0.0052	0.0100	0.0055

Sequence No.

	21	22	23	24	25	26	27	28	29	30
$D_N = \max\lvert F_N(X) - F_0(X)\rvert$	0.0067	0.0074	0.0062	0.0095	0.0069	0.0059	0.0079	0.0104	0.0055	0.0059

Sequence No.

	31	32	33	34	35	36	37	38	39	40
$D_N = \max\lvert F_N(X) - F_0(X)\rvert$	0.0058	0.0059	0.0147*	0.0114	0.0060	0.0051	0.0071	0.0054	0.0192*	0.0067

Sequence No.

	41	42	43	44	45	46	47	48	49	50
$D_N = \max\lvert F_N(X) - F_0(X)\rvert$	0.0117	0.0108	0.0043	0.0046	0.0076	0.0050	0.0079	0.0107	0.0197*	0.0066

The critical value of D_K for $N = 10{,}000$ and the 0.05 probability level is 0.0136.

TABLE 5.22

RESULTS OF THE KOLMOGOROV-SMIRNOV GOODNESS-OF-FIT TEST FOR UNIFORMITY OF 100 SEQUENCES GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Concluded

Sequence No.

	51	52	53	54	55	56	57	58	59	60
$D_N = \max\lvert F_N(X) - F_0(X)\rvert$	0.0042	0.0066	0.0057	0.0060	0.0091	0.0105	0.0062	0.0081	0.0083	0.0057

Sequence No.

	61	62	63	64	65	66	67	68	69	70
$D_N = \max\lvert F_N(X) - F_0(X)\rvert$	0.0064	0.0062	0.0037	0.0107	0.0103	0.0041	0.0089	0.0060	0.0079	0.0096

Sequence No.

	71	72	73	74	75	76	77	78	79	80
$D_N = \max\lvert F_N(X) - F_0(X)\rvert$	0.0056	0.0104	0.0071	0.0102	0.0032	0.0068	0.0048	0.0069	0.0079	0.0073

Sequence No.

	81	82	83	84	85	86	87	88	89	90
$D_N = \max\lvert F_N(X) - F_0(X)\rvert$	0.0066	0.0066	0.0068	0.0079	0.0064	0.0076	0.0105	0.0067	0.0095	0.0096

Sequence No.

	91	92	93	94	95	96	97	98	99	100
$D_N = \max\lvert F_N(X) - F_0(X)\rvert$	0.0061	0.0114	0.0075	0.0069	0.0058	0.0105	0.0075	0.0086	0.0099	0.0072

The critical value of D_k for $N = 10{,}000$ and the 0.05 probability level is 0.0136.

TABLE 5.30

RESULTS OF A TEST ON THE MEANS OF 50 SEQUENCES GENERATED BY EQUATION (2.2.1)

Theoretical mean $\mu = 0.00000$

Sequence No.

	1	2	3	4	5	6	7	8	9	10
Observed mean $\bar{x}$	0.00596	0.00037	-0.00916	0.00681	-0.00187	0.01150	-0.00456	0.00986	-0.01137	-0.00442
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	0.596	0.037	-0.916	0.681	-0.188	1.150	-0.456	0.986	-1.137	-0.442

Sequence No.

	11	12	13	14	15	16	17	18	19	20
Observed mean $\bar{x}$	0.00352	0.00565	-0.00259	-0.00886	-0.00306	-0.00126	0.01531	0.00130	-0.00888	-0.00623
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	0.351	0.565	-0.259	-0.886	-0.306	-0.126	1.531	0.130	-0.887	-0.623

Sequence No.

	21	22	23	24	25	26	27	28	29	30
Observed mean $\bar{x}$	0.01529	-0.00095	-0.00133	0.00630	0.00220	0.00203	0.00697	0.00677	0.00162	0.00932
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	1.528	-0.096	-0.133	0.630	0.220	0.203	0.697	0.677	0.162	0.932

Sequence No.

	31	32	33	34	35	36	37	38	39	40
Observed mean $\bar{x}$	0.00791	0.00848	-0.00829	-0.00598	-0.01329	-0.00177	-0.00238	0.00011	-0.00166	-0.00309
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	0.791	0.848	-0.829	-0.597	-1.329	-0.177	-0.238	0.011	-0.166	-0.309

Sequence No.

	41	42	43	44	45	46	47	48	49	50
Observed mean $\bar{x}$	0.01063	0.00440	-0.00137	-0.00499	-0.01428	-0.00839	0.00849	-0.00275	-0.00898	0.00273
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	1.063	0.441	-0.137	-0.499	-1.428	-0.842	0.849	-0.275	-0.898	0.273

The critical values of Z for the 0.05 probability level are ± 1.96.

TABLE 5.31

RESULTS OF A TEST ON THE MEANS OF 50 SEQUENCES GENERATED BY EQUATION (2.2.2)

Theoretical mean μ = 0.00000

Sequence No.

	1	2	3	4	5	6	7	8	9	10
Observed mean $\overline{x}$	0.00021	-0.00803	0.00009	-0.00913	0.01333	-0.00105	-0.00051	0.01258	0.00041	0.00719
$Z = \frac{\overline{x} - \mu}{\sigma/\sqrt{N}}$	0.021	-0.803	0.009	-0.913	1.333	-0.105	-0.051	1.258	0.041	0.719

Sequence No.

	11	12	13	14	15	16	17	18	19	20
Observed mean $\overline{x}$	0.00303	-0.00354	-0.00584	-0.01106	-0.01187	-0.00859	-0.00900	0.00640	0.01646	-0.00672
$Z = \frac{\overline{x} - \mu}{\sigma/\sqrt{N}}$	0.303	-0.354	-0.584	-1.106	-1.187	-0.859	-0.900	0.640	1.646	-0.672

Sequence No.

	21	22	23	24	25	26	27	28	29	30
Observed mean $\overline{x}$	0.00623	0.00460	-0.01609	-0.00597	-0.00498	0.00839	-0.00532	-0.00575	-0.00033	0.00781
$Z = \frac{\overline{x} - \mu}{\sigma/\sqrt{N}}$	0.623	0.460	-1.609	-0.606	-0.498	0.839	-0.532	-0.575	-0.033	0.781

Sequence No.

	31	32	33	34	35	36	37	38	39	40
Observed mean $\overline{x}$	-0.01114	-0.01479	0.00290	0.00476	0.00498	-0.00145	-0.00851	-0.00154	0.00190	-0.00004
$Z = \frac{\overline{x} - \mu}{\sigma/\sqrt{N}}$	-1.114	-1.479	0.290	0.476	0.498	-0.145	-0.851	-0.154	0.190	-0.004

Sequence No.

	41	42	43	44	45	46	47	48	49	50
Observed mean $\overline{x}$	0.00473	0.00082	-0.01424	-0.00019	-0.01523	0.00008	-0.00197	-0.00872	-0.01255	-0.00444
$Z = \frac{\overline{x} - \mu}{\sigma/\sqrt{N}}$	0.473	0.082	-1.424	-0.019	-1.523	0.008	-0.197	-0.872	-1.255	-0.444

The critical values of Z for the 0.05 probability level are ± 1.96.

TABLE 5.32

RESULTS OF A TEST ON THE MEANS OF 100 SEQUENCES GENERATED BY

THE UNIFORM NUMBER GENERATOR (2.1.4)

Theoretical mean $\mu = 0.50000$

Sequence No.

	1	2	3	4	5	6	7	8	9	10
Observed mean $\bar{x}$	0.49981	0.50076	0.49780	0.50183	0.49904	0.50899	0.50346	0.49878	0.50226	0.50601
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	-0.066	0.263	-0.762	0.632	-0.331	3.138*	1.202	-0.424	0.782	2.100*

Sequence No.

	11	12	13	14	15	16	17	18	19	20
Observed mean $\bar{x}$	0.50238	0.50086	0.50295	0.50033	0.50417	0.50145	0.49802	0.49979	0.49493	0.50089
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	0.825	0.298	1.022	0.113	1.452	0.500	-0.683	-0.072	-1.756	0.307

Sequence No.

	21	22	23	24	25	26	27	28	29	30
Observed mean $\bar{x}$	0.50252	0.50195	0.50225	0.50523	0.49780	0.50084	0.50450	0.49732	0.50182	0.49893
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	0.879	0.675	0.780	1.822	-0.766	0.293	1.569	-0.920	0.628	-0.368

Sequence No.

	31	32	33	34	35	36	37	38	39	40
Observed mean $\bar{x}$	0.49936	0.50085	0.50809	0.50417	0.49954	0.49813	0.50313	0.50166	0.50870	0.50261
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	-0.219	0.295	2.829*	1.453	-0.158	-0.649	1.086	0.575	3.034*	0.907

Sequence No.

	41	42	43	44	45	46	47	48	49	50
Observed mean $\bar{x}$	0.50567	0.50449	0.49968	0.50048	0.50281	0.50026	0.50296	0.50559	0.50908	0.50227
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	1.980*	1.564	-0.111	0.168	0.976	0.092	1.025	1.952	3.173*	0.785

Critical values of Z for the 0.05 probability level are ±1.96.

TABLE 5.32

RESULTS OF A TEST ON THE MEANS OF 100 SEQUENCES GENERATED BY

THE UNIFORM NUMBER GENERATOR (2.1.4) - Concluded

Theoretical mean $\mu = 0.50000$

Sequence No.

	51	52	53	54	55	56	57	58	59	60
Observed mean $\bar{x}$	0.49879	0.50199	0.50167	0.49995	0.49658	0.50550	0.50152	0.49749	0.49600	0.49867
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	-0.419	0.689	0.578	-0.017	-1.173	1.916	0.529	-0.861	-1.387	-0.459

Sequence No.

	61	62	63	64	65	66	67	68	69	70
Observed mean $\bar{x}$	0.50094	0.50190	0.50030	0.50442	0.50482	0.50046	0.50483	0.49791	0.50173	0.50518
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	0.328	0.658	0.105	1.539	1.677	0.159	1.677	-0.725	0.596	1.785

Sequence No.

	71	72	73	74	75	76	77	78	79	80
Observed mean $\bar{x}$	0.49854	0.50478	0.49836	0.50463	0.49963	0.49865	0.50060	0.50280	0.50286	0.50267
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	-0.504	1.656	-0.569	1.610	-0.129	-0.466	0.207	0.972	0.991	0.923

Sequence No.

	81	82	83	84	85	86	87	88	89	90
Observed mean $\bar{x}$	0.50238	0.50246	0.50243	0.49836	0.50229	0.50317	0.50403	0.50242	0.50381	0.50419
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	0.829	0.856	0.839	-0.564	0.792	1.100	1.405	0.840	1.319	1.453

Sequence No.

	91	92	93	94	95	96	97	98	99	100
Observed mean $\bar{x}$	0.50210	0.50576	0.50208	0.49696	0.50241	0.50445	0.50319	0.49750	0.49513	0.50185
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$	0.726	2.003*	0.724	-1.056	0.838	1.541	1.107	-0.863	-1.691	0.640

Critical values of Z for the 0.05 probability level are ±1.96.

TABLE 5.33

RESULTS OF A TEST ON THE VARIANCES OF 50 SEQUENCES GENERATED BY EQUATION (2.2.1)

Theoretical variance $\sigma^2 = 1.00000$

Sequence No.

	1	2	3	4	5	6	7	8	9	10
Observed variance s^2	1.01487	0.98595	1.00221	1.00878	0.99124	1.00398	0.96974	1.02175	0.99535	1.00573
$\chi^2 = \frac{(N-1)s^2}{\sigma^2}$	10,147.68	9858.51	10,021.09	10,086.79	9911.40	10,038.79	9696.43*	10,216.48	9952.50	10,056.29

Sequence No.

	11	12	13	14	15	16	17	18	19	20
Observed variance s^2	1.01894	1.00200	0.95713	0.98226	1.00070	1.01665	0.97410	0.98913	0.97495	1.00025
$\chi^2 = \frac{(N-1)s^2}{\sigma^2}$	10,188.38	10,019.00	9570.34*	9821.64	10,006.00	10,165.48	9740.03	9890.31	9748.53	10,001.50

Sequence No.

	21	22	23	24	25	26	27	28	29	30
Observed variance s^2	0.96062	0.97339	0.99693	0.99015	0.99662	0.97182	0.98536	0.98801	1.00273	0.97883
$\chi^2 = \frac{(N-1)s^2}{\sigma^2}$	9605.24*	9732.93	9968.30	9900.51	9965.20	9717.23*	9852.61	9879.11	10,026.30	9787.32

Sequence No.

	31	32	33	34	35	36	37	38	39	40
Observed variance s^2	0.98439	0.99557	0.97924	1.00680	0.97745	0.99138	1.01921	1.01680	1.01005	0.99281
$\chi^2 = \frac{(N-1)s^2}{\sigma^2}$	9842.92	9954.70	9791.42	10,067.09	9773.52	9912.81	10,191.08	10,166.98	10,099.49	9927.11

Sequence No.

	41	42	43	44	45	46	47	48	49	50
Observed variance s^2	0.96364	0.99450	0.98150	0.96468	0.98400	0.99270	1.01361	1.02164	0.99776	1.00412
$\chi^2 = \frac{(N-1)s^2}{\sigma^2}$	9635.44*	9944.01	9814.02	9645.84*	9839.02	9926.01	10,135.09	10,215.38	9976.60	10,040.20

Critical values for χ^2 with 9999 degrees of freedom and the 0.05 probability level are: $\chi^2_{.975} = 9723.15$; $\chi^2_{.025} = 10,277.48$.

TABLE 5.34

RESULTS OF A TEST ON THE VARIANCES OF 50 SEQUENCES GENERATED BY EQUATION (2.2.2)

Theoretical variance $\sigma^2 = 1.00000$

Sequence No.

	1	2	3	4	5	6	7	8	9	10
Observed variance s^2	0.99181	0.99725	1.00645	0.99772	1.00870	0.95544	0.97748	1.00697	0.99018	0.94887
$\chi^2 = \frac{(N-1)s^2}{\sigma^2}$	9917.10	9971.51	10,063.50	9976.20	10,086.00	9553.40*	9773.81	10,068.70	9900.81	9487.71*

Sequence No.

	11	12	13	14	15	16	17	18	19	20
Observed variance s^2	0.99380	0.99064	0.97703	1.00265	0.97562	0.98843	0.99545	0.99629	1.01247	0.98584
$\chi^2 = \frac{(N-1)s^2}{\sigma^2}$	9937.00	9905.40	9769.31	10,025.50	9755.20	9883.30	9953.51	9961.90	10,123.70	9857.41

Sequence No.

	21	22	23	24	25	26	27	28	29	30
Observed variance s^2	0.99048	0.98388	1.01708	0.97095	1.02144	0.97319	1.01920	0.98408	1.00199	0.99279
$\chi^2 = \frac{(N-1)s^2}{\sigma^2}$	9903.80	9837.80	10,169.81	9708.50*	10,213.40	9730.90	10,191.00	9839.80	10,018.90	9926.90

Sequence No.

	31	32	33	34	35	36	37	38	39	40
Observed variance s^2	1.00819	1.00165	1.01097	1.02662	1.00845	0.99098	0.98147	0.97163	0.97338	0.98282
$\chi^2 = \frac{(N-1)s^2}{\sigma^2}$	10,080.90	10,015.50	10,108.70	10,265.20	10,083.51	9908.80	9813.70	9715.30*	9732.81	9827.21

Sequence No.

	41	42	43	44	45	46	47	48	49	50
Observed variance s^2	0.97315	1.01632	0.99389	0.96750	0.98152	1.00766	0.97579	0.99098	1.00229	0.98571
$\chi^2 = \frac{(N-1)s^2}{\sigma^2}$	9730.50	10,162.20	9937.91	9674.01*	9814.20	10,075.61	9756.90	9908.80	10,021.90	9856.11

Critical values for χ^2 with 9999 degrees of freedom and the 0.05 probability level are: $\chi^2_{.975} = 9723.15$; $\chi^2_{.025} = 10,277.48$.

TABLE 5.35

RESULTS OF A TEST ON THE STANDARD DEVIATION OF 100 SEQUENCES GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4)

Theoretical variance $\sigma^2 = 0.08333$

Sequence No.

	1	2	3	4	5	6	7	8	9	10
Observed variance s^2	0.08374	0.08304	0.08235	0.08336	0.08359	0.08201	0.08286	0.08312	0.08394	0.08198
$Z = \frac{s - \sigma}{\sigma/\sqrt{2N}}$	0.350	-0.253	-0.839	0.026	0.221	-1.125	-0.404	-0.180	0.516	-1.150

Sequence No.

	11	12	13	14	15	16	17	18	19	20
Observed variance s^2	0.08308	0.08397	0.08338	0.08343	0.08266	0.08381	0.08401	0.08410	0.08291	0.08307
$Z = \frac{s - \sigma}{\sigma/\sqrt{2N}}$	-0.212	0.542	0.037	0.080	-0.572	0.402	0.576	0.645	-0.356	-0.227

Sequence No.

	21	22	23	24	25	26	27	28	29	30
Observed variance s^2	0.08253	0.08309	0.08350	0.08234	0.08235	0.08339	0.08243	0.08489	0.08379	0.08376
$Z = \frac{s - \sigma}{\sigma/\sqrt{2N}}$	-0.683	-0.204	0.140	-0.848	-0.839	0.046	-0.771	1.313	0.384	0.363

Sequence No.

	31	32	33	34	35	36	37	38	39	40
Observed variance s^2	0.08468	0.08308	0.08184	0.08215	0.08405	0.08329	0.08288	0.08309	0.08234	0.08309
$Z = \frac{s - \sigma}{\sigma/\sqrt{2N}}$	1.136	-0.217	-1.269	-1.007	0.607	-0.039	-0.384	-0.203	-0.848	-0.203

Sequence No.

	41	42	43	44	45	46	47	48	49	50
Observed variance s^2	0.08213	0.08240	0.08339	0.08290	0.08319	0.08324	0.08318	0.08207	0.08198	0.08348
$Z = \frac{s - \sigma}{\sigma/\sqrt{2N}}$	-1.027	-0.792	0.048	-0.368	-0.124	-0.081	-0.133	-1.073	-1.156	0.126

Critical values of Z for the 0.05 probability level are ±1.96.

TABLE 5.35

RESULTS OF A TEST ON THE STANDARD DEVIATION OF 100 SEQUENCES GENERATED BY THE UNIFORM NUMBER GENERATOR (2.1.4) - Concluded

Theoretical variance $\sigma^2 = 0.08333$

Sequence No.

	51	52	53	54	55	56	57	58	59	60
Observed variance s^2	0.08307	0.08268	0.08306	0.08258	0.08492	0.08225	0.08281	0.08494	0.08303	0.08388
$Z = \frac{s - \sigma}{\sigma/\sqrt{2N}}$	-0.226	-0.559	-0.229	-0.641	1.337	-0.923	-0.442	1.353	-0.254	0.467

Sequence No.

	61	62	63	64	65	66	67	68	69	70
Observed variance s^2	0.08324	0.08309	0.08310	0.08246	0.08278	0.08300	0.08287	0.08327	0.08437	0.08232
$Z = \frac{s - \sigma}{\sigma/\sqrt{2N}}$	-0.080	-0.209	-0.197	-0.739	-0.473	-0.283	-0.391	-0.052	0.876	-0.864

Sequence No.

	71	72	73	74	75	76	77	78	79	80
Observed variance s^2	0.08436	0.08334	0.08320	0.08283	0.08338	0.08366	0.08322	0.08305	0.08317	0.08335
$Z = \frac{s - \sigma}{\sigma/\sqrt{2N}}$	0.873	0.003	-0.114	-0.430	0.042	0.273	-0.097	-0.237	-0.139	0.010

Sequence No.

	81	82	83	84	85	86	87	88	89	90
Observed variance s^2	0.08283	0.08233	0.08346	0.08423	0.08384	0.08288	0.08226	0.08282	0.08317	0.08297
$Z = \frac{s - \sigma}{\sigma/\sqrt{2N}}$	-0.430	-0.854	0.104	0.760	0.426	-0.384	-0.892	-0.433	-0.138	-0.310

Sequence No.

	91	92	93	94	95	96	97	98	99	100
Observed variance s^2	0.08358	0.08277	0.08254	0.08300	0.08268	0.08339	0.08285	0.08385	0.08302	0.08317
$Z = \frac{s - \sigma}{\sigma/\sqrt{2N}}$	0.207	-0.483	-0.673	-0.282	-0.558	0.048	-0.408	0.439	-0.269	-0.139

Critical values of Z for the 0.05 probability level are ±1.96.

VI. SUMMARY

This study investigated a generator producing numbers assumed to be random and uniformly distributed on the internal (0,1), and two equations each producing numbers assumed to be random and normally distributed with zero mean and unit variance.

A total of 200 sequences each with 10,000 observations were generated for this study. Of these, 100 were generated by the uniform number generator (2.1.4), 50 by equation (2.2.1), and the remaining 50 by equation (2.2.2). A computer program was written by the author for this purpose.

Based on the results of the various statistical tests performed on all sequences, the following conclusions may be drawn:

(1) It is apparent that the uniform random number generator (2.1.4) is generating numbers that seem to be random and uniformly distributed on the unit interval. Hence, there is an indication that this generator should be a useful addition to the state-of-the-art.

(2) It is reasonable to believe that equation (2.2.1) generates numbers that appear to be random and normally distributed with zero mean and unit variance.

(3) Although equation (2.2.2) generates numbers that appear to be normally distributed with zero mean and unit variance, the existence of nonrandomness among the numbers it generates is apparent based on the results of the test for zero correlation and the test for runs above and below the mean of the specified distribution function. Hence, this study

tends to conclude that equation (2.2.1) is the better normal random number generator.

ACKNOWLEDGEMENTS

The author would like to express his deep appreciation and indebtedness to his thesis advisor, Dr. Richard G. Krutchkoff, for his patience and prudent guidance throughout this study.

Appreciation is also expressed to Mr. Whitney Johnson for serving as coreader and for his constructive criticism.

The author also wishes to express his deep appreciation to Dr. Boyd Harshbarger and the Department of Statistics for making possible a teaching assistantship during part of my residency. Appreciation is also expressed to all members of the Department of Statistics for their help and guidance.

Deep appreciation is expressed to the National Aeronautics and Space Administration and especially to the Analysis and Computation Division, Langley Research Center, for financial support during two-thirds of my residency and for making this study possible.

Kind regards and thanks are extended to Mr. Richard D. Hofler for his valuable assistance in the computer program.

BIBLIOGRAPHY

1. Tausworthe, R. C.: Random Numbers Generated by Linear Recurrence Modulo Two, Mathematics of Computation, Vol. 19, 1965, pp. 201-209.

2. Golomb, S. W.: Sequences with Randomness Properties, Martin Co., Baltimore, Md., 1955.

3. Zierler, Neal: Linear Recurring Sequences, J. Soc. Indust. Appl. Math., Vol. 7, 1959, pp. 31-48.

4. Baumert, L., et al.: Coding Theory and its Applications to Communications Systems, Report 32-167, Jet Propulsion Laboratory, Pasadena, Calif., 1961.

5. Watson, E. J.: Primitive Polynomials (Mod 2), Mathematics of Computation, Vol. 16, 1962, pp. 368-369.

6. Box, G. E. P.; and Muller, M. G.: A Note on the Generation of Random Normal Deviates, The Annals of Mathematical Statistics, Vol. 29, 1958, pp. 610-611.

7. Mood, A. M.; and Graybill, F. A.: Introduction to the Theory of Statistics, McGraw-Hill Book Company, Inc., New York, 1963, pp. 409-412.

8. Gorenstein, S.: Random Number Generation for the General Purpose Systems Simulator, Technical Report 17-161, IBM, Advanced Systems Development Division, Yorktown Heights, N. Y., 1966.

9. Wald, A.; and Wolfowitz, J.: An Exact Test for Randomness in the Non-Parametric Case Based on Serial Correlation, Annals of Mathematical Statistics, Vol. 14, pp. 378-388.

10. Miller, I.; and Freund, J. E.: Probability and Statistics for Engineers, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1965, pp. 219-221.

11. Massey, F. J. Jr.: The Kolmogorov-Smirnov Test for Goodness of Fit, Journal of the American Statistical Association, Vol. 46, 1951, pp. 68-78.

12. Hoel, P. G.: Introduction to Mathematical Statistics, 3rd ed., John Wiley and Sons, Inc., New York, 1965, pp. 335-341.

13. Kolmogorov, A.: Sulla Determinazione Empirica di una Legge di Distributione, Geornale dell'Istituto Italiano degli Attuari, 4, 1933, pp. 1-11.

14. Smirnov, H.: Sur les Ecarts de la Courbe de Distribution Empirique, Recueil Mathematique, N.S. 6, 1939, pp. 3-26.

15. Massey, F. J., Jr.: A Note on the Power of a Non-Parametric Test, Annals of Mathematical Statistics, Vol. 21, 1950, pp. 440-443.

16. Pearson, K.: On a Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it Can be Reasonably Supposed to Have Arisen in Random Sampling, Phil. Mag., (5), 50, 157.

17. Kendall, M. G.; and Stuart, A.: The Advanced Theory of Statistics, Vol. II, Hafner Publishing Co., New York, 1961, pp. 420-441.

18. Mann, H. B.; and Wald, A.: On the Choice of the Number of Intervals in the Application of the Chi-square Test, Annals of Mathematical Statistics, 13, 1942, pp. 306-317.

19. Williams, C. A., Jr.: On the Choice of the Number and Width of Classes for the Chi-square Test for Goodness of Fit, Journal of the American Statistical Association, 45, 1950, pp. 77-86.

20. Stockmal, Frank: Calculations with Pseudo-Random Numbers, Journal of the Association for Computing Machinery, 11, 1964, pp. 41-52.

VITA

The author was born in Karpenissi, Greece on April 23, 1939. He received his Bachelor of Science degree in Mathematics in June, 1963, from Virginia Polytechnic Institute.

After his graduation, he accepted a position as an Aerospace Technologist with the National Aeronautics and Space Administration, Langley Research Center, where he has been employed up to the present. In September, 1964, he enrolled at Virginia Polytechnic Institute as a graduate student and since has been working toward the Master of Science degree in Statistics.